

斯蒂格利茨经济学文集

第二卷

中国金融出版社

斯蒂格利茨经济学文集

Selected Works on Economics of Joseph E. Stiglitz

第一卷
信息经济学：基本原理
Vol.1 Economics of Information: Basic Principles

第二卷
信息经济学：应用
Vol.2 Economics of Information: Applications

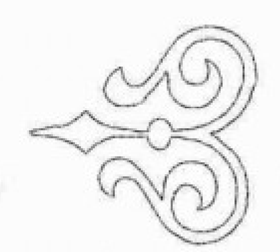
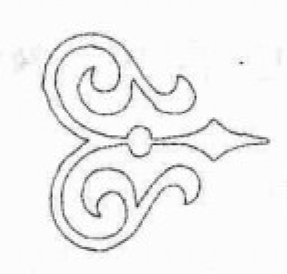
第三卷
微观经济学：不确定性与研发
Vol.3 Micro Economics, Including Uncertainty, R&D

第四卷
宏观经济学：增长与分配
Vol.4 Macro Economics, Growth and Distribution

第五卷
公共财政
Vol.5 Public Finance

第六卷
发展与发展政策
Vol.6 Development and Development Policy

斯蒂格利茨经济学文集


第二卷


信息经济学：应用

Economics of Information: Applications

［美］约瑟夫·斯蒂格利茨　著

纪沫　陈佳　刘海燕　译

中国金融出版社

总 策 划：林铁钢
责任编辑：何　为
责任校对：孙　蕊
责任印制：裴　刚

图书在版编目（CIP）数据

信息经济学：应用（Xinxi Jingjixue：Yingyong）/（美）斯蒂格利茨（Stiglitz，J. E.）著；纪沫，陈佳，刘海燕译．—北京：中国金融出版社，2007.7
（斯蒂格利茨经济学文集；2）
书名原文：Economics of Information：Applications
ISBN 978-7-5049-4250-0

Ⅰ．信…　Ⅱ．①斯…②纪…③陈…④刘…　Ⅲ．信息经济学—文集
Ⅳ．F062.5-53

中国版本图书馆 CIP 数据核字（2007）第 111882 号

出版发行　中国金融出版社
社址　北京市广安门外小红庙南里 3 号
市场开发部　（010）63272190，66070804（传真）
网上书店　http://www.chinafph.com
（010）63286832，63365686（传真）
读者服务部　（010）66070833，82672183
邮编　100055
经销　新华书店
印刷　北京汇林印务有限公司
尺寸　155 毫米×235 毫米
印张　22
字数　395 千
版次　2007 年 9 月第 1 版
印次　2007 年 9 月第 1 次印刷
定价　680.00 元（1—6 卷）
ISBN 978-7-5049-4250-0/F.3810

目 录

第二卷导言

第一卷详细阐述了信息经济学的基本原理，本卷进一步考察其在不同问题上的应用。资本市场、劳动市场和产品市场等每个市场都会受到不完美信息的影响。不完美信息会影响企业和家庭的行为。不完美信息会改变市场均衡的本质。包括规制和私有化在内的政府政策也会受到不完美信息的根本性影响。

收录于此卷中的论文奠定了新企业理论的基础（第三卷进一步阐述新企业理论，而新企业理论对宏观经济学的影响则是第四卷的中心议题）。旧企业理论的内容并不丰富：企业只是采用给定的生产函数（即投入与产出之间的关系），并选择投入和产出去使利润（市场价值）最大化。企业不会想到去激励工人——只要工人完成了约定的工作任务，就能够获得约定的报酬；如果没有完成约定的工作任务，那么就不能得到约定的报酬。企业不会考虑融资的问题——它们会以现成的利率贷到资金，直到资本的边际收益等于边际成本为止。当经济学家开始讨论企业财务结构问题时——是通过债权还是股权融资——（在莫迪利安尼和米勒的开创性论文中）得出的结论是，企业的财务结构无关紧要。

本卷的第一编讨论了决策制定的组织中的基本问题。这是标准新古典模型（旧范式）不会去探讨的另一个问题。存在不完美信息，就意味着个体可能会犯错误——人人难免有错。我研究了一个简单而重要的问题：组织设计影响组织犯错误的概率的机制是什么？如果组织设计会对组织犯错误的概率产生影响的话，那么它影响的到底是哪一类错误？长期以来，人们都认为，为了尽量降低个体可能会犯的错误（或者个体的恶意行为）带来的损失，就应该采取相互制衡的方法。我们在一个简单的背景下，即组织的任务是决定是否开展某个项目，分析了组织设计对组织犯错误的概率的影响。分析结果表明，与等级制企业相比，分权的组织（即我们通常所说的多头的组织）开展一个坏项目的可能性更大。不过，分权的组织开展好项目的可能性也更大。这就意味着，分权的组织形式有着明显的优势。当社会缺乏好

点子的时候，分权的组织这种优势就体现得更加明显。此项研究的数理基础和思想观念已经得到了广泛的运用和与推广。[1]

第二编提供了将新的信息范式应用于资本市场的三个例子。第一篇论文分析了两个基本问题。长期以来，经济学家都没有对公司财务问题给予足够的重视——它只是商学院讲授的一门课程而已，而且也不太重要。莫迪利安尼和米勒是最先深入思考公司财务问题的经济学家，他们的研究结果表明忽视公司财务问题事实上是正确的。这是因为，他们指出公司财务结构并不重要。这个结论非比寻常：如果它是对的，那它就为市场的非理性提供了不可辩驳的证据。因为华尔街大部分人的工作都与协助各个公司管理其财务有关，而且华尔街上这些人似乎也认为公司财务是非常重要的。我们的分析结果显示，由于公司财务结构能传递信息，所以它是非常重要的：公司所有者将其持有的股票出售给外部人这个事实表明，外部人也许高估了公司股票的价值。公司所有者不愿意卖掉其手中的股票，就相当于向外部人发送了一个信号，即他们对公司的发展有信心。公司财务结构实际上就成了实现甄别/信号发送均衡的机制。[2]

此文的第二个研究结论提出了一个根本性的问题：为什么理性的个体会在股票市场上进行交易？他们买卖股票不会比只购买市场组合要更好，当然也不会比只购买市场组合要更差。如果有人买卖股票的收益高于市场收益，那么肯定会有人买卖股票的收益低于市场收益。假如后者意识到这一点，那他们就可以转而购买市场组合。但是，如果所有不知情的或能力不强的市场参与人都采取这种策略的话，那么哪些人才是赚到钱的知情人和能力强的人？当然，从理论上讲，市场可以将证券与交易者“匹配”，从而让他们为其无法直接投保的人力资本或其他资本进行保险，至少是进行部分保险。但是，股票市场上的活动很少是这类匹配活动。简而言之，股票市场在很大程

1　例如，在 Sah 和 Stiglitz（1988a）这篇文章中，我们分析了决策委员会的最优设计中更具有一般性的问题——委员会的规模以及决策规则（比如在由 n 个成员组成的委员会中，需要有 k 个成员的同意才能做出一项决议）。在 Sah 和 Stiglitz（1988b）这篇文章中，我们奠定了此一般性理论的数理基础。在 Sah 和 Stiglitz（1991）这篇论文中，我们将此理论应用于一个特别的情形之中——管理者选择问题，即每一任管理者如何挑选其继任者。我们的分析再一次表明，分散的决策能使公司绩效的波动幅度变得更小。Stiglitz（1991）评述了这一领域中（在 90 年代初看来）的关键性议题。此外，Stiglitz（1989）还提出了组织设计研究领域中许多重要议题，其中就包括协调问题。这些议题与 Stiglitz（1991）评述的议题有相当大的差异。

2　Ross 以及 Leland 和 Pyle 同时也得到了类似的结果。Greenwald，Stiglitz 和 Weiss（1984）（重印于第一卷之中）以及 Majluf 和 Myers（1984）所做的数理分析与此略有不同。

度上就是有钱人的赌场。如果这个观点是正确的——它一直都没有受到真正的挑战——那么它对我们认识股票市场在市场经济中的作用就有重要含义。[3]

第二篇文章继续考察信贷配给的作用。传统的经济学一直都假定各人都能以（经过风险调整后的）市场利率借到他们需要的资金。事实上，许多人的借款行为面临着非常严重的数量限制。Stiglitz 和 Weiss（1981）[4] 对此现象给出了一个解释：利率越高，贷款人就要与（违约概率更高的）质量更差的贷款申请人打交道；而且更高的利率就会引致借款人开展更有风险的项目。由于存在这些逆向选择效应和激励效应，所以提高利率实际上可能会降低预期收益。而且如果实际情况确实如此，那么可能就会存在一个带有信贷配给的竞争均衡。那些无法得到贷款的潜在的借款人，既不会去银行借款，也不能支付更高的利率：贷款人假定收回贷款的机会很少，以至于将贷款借给这样的借款人没有任何意义。我们指出，信贷配给还有更深层次的原因：假如借款人知道，如果他们在某一期表现不佳，那么在以后各期他们都无法得到贷款，则借款人表现良好（不去承担过度的风险）的激励就会得到增强。在信贷和劳动市场上，终止合约是有效的激励机制——比其他像收取更高的利率这样的威胁都有效。此文的分析还具有更深层次的重要含义：各期的市场是相互联系的；让同一个贷款人在不同的期间贷款给借款人是更有效率的安排；而且各期市场之间的相互联系，削弱了市场的有效竞争。使市场变得有竞争性的一个因素就是“转换”，雇员转换为雇主，或借款人转换成贷款人；但是，当各期市场之间存在相互关联的关系时，这种转换就变得更加困难了。此外，转换行为本身也会削弱作为一种激励机制的终止合约这种威胁的有效性，因此转换行为本身也会对市场效率产生负面影响。

这一编的最后一篇文章既对带有信息问题的市场理论做出了奠基性贡献，又具有重大的现实意义。它为各种小额信贷计划奠定了理论基础。小额信贷计划是由孟加拉国的 Grameen 银行创造出来的，穆哈默德·约努斯（Mohammed Yunus）也因此而获得了2006年度诺贝尔和平奖。这些小额信贷计划使得最贫穷的人也能借到贷款。尤为引人注目的是，这些小额贷款计划的贷款偿付比率还很高。在信息经济学领域，我早期的研究就是去考察提

3　有必要将这篇与 Grossman 和 Stiglitz（1976，1980）这两篇文章结合起来阅读。后两篇文章分析了资本市场可能是无法实现信息效率的原因。近来，人们逐渐意识到资本市场易受到非理性的乐观主义和悲观主义的影响。而且这些非理性因素会对宏观经济产生显著的影响。最近的研究也探讨了在不完美信息条件下，为什么会出现理性的羊群行为。

4　此文重印于卷一之中。

高偿付比率的两个机制——（正如前一篇论文所讨论的）激励和监督。不过，监督通常都是很困难的事情，而且成本也很高。过去，监督都是层级式监督，例如老板监督工人、银行监督借款人等等。在《同侪监督与信贷市场》这篇文章中，我提出了一种替代性的监督机制，即由同侪来进行监督。工人可以相互监督，或者是借款人相互监督。当然，要让同侪监督发挥作用，该群体就必须要有积极性去实施监督，这篇论文的分析指出了怎么样才能做到这一点。[5]

在经济学中，没有其他问题比激励更加重要了。事实上，经济学有时就被称为激励的科学。经济学家经常会提到激励的作用，不过奇怪的是，在信息经济学诞生之前，经济学家很少会谈到激励。在标准的新古典模型中，每个人只要完成了合同约定的任务，他就能够获得其报酬，否则就得不到报酬。企业不需要去激励工人。然而，激励问题是企业面临的核心问题。

可以用一个简单的方法去激励工人努力工作：例如，以计件工资为基础，根据工人的产出或者绩效支付其报酬。我要问的第一个问题是，为什么要这样做?《分成租佃制中的激励与风险分担》[6]这篇文章从风险与信息的角度回答了这个问题。在投入（工人的努力）与产出之间并不存在简单的对应关系（气候也会影响产出），而且很难从可以观察的产出中推测工人的投入。因此，以产出为基础来支付报酬，就会使工人承担相当大的风险。以投入为基础来支付报酬也许更好，但是投入通常都是不可以观察的。雇主可以根据工人的劳动时间来支付工人工资，并且严格监督他们以确保工人努力工作。这样做的成本很高，而且也不是非常有效。工人可以租赁土地；由于他们可以获得其劳动投入带来的所有收益，因此他们努力工作的积极性就会非常高；但是，此时他们又需要承担大量的风险，因为即便是天气不好以及农作物收成不高时，他们仍然需要支付租金。我的研究表明，分成租佃制实际上是在风险分担和有效激励之间做出的一种折中的选择。实际上是可以设计出最优的激励机制来权衡风险分担与有效激励。此卷的第三编第一篇文章《激励、风险和信息：对等级理论的评论》最重要的一个贡献就是，建立了数理模型去分析最优激励机制。还有其他许多原因使得企业无法过多地采取计件工资制。当许多人在一个团队中工作时，即便不是不可能，那也很难区分

5 我和阿诺特合作的论文分析了其他监督机制，参见 Arnott and Stiglitz，1991，"Moral Hazard and Non-Market Institutions：Dysfunctional Crowding Out or Peer Monitoring，" American Economic Review，Vol. 81（1），pp. 179-190。

6 Review of Economic Studies，41(2)，April 1974，pp. 219-255，重印于第一卷之中。

每个人的贡献。更一般地讲，组织结构会影响其进行监督和提供激励的能力。

组织设计与激励之间的关系，是我在《奖赏与激励》这篇文章中继续探讨的主题。许多企业对那些表现最好的工人进行激励，或者是惩罚（解雇）那些表现最差的工人。这或多或少会使得工人相互之间进行竞争。为什么这样说？为什么要使用相对绩效作为支付工人工资的基础，而不是使用绝对绩效？毫不奇怪，答案又是不完全信息。雇主通常都不知道工作任务到底又多难？如果他知道工作任务的难度，那么他就能更好地“微调”激励性报酬。但是，他可以通过观察所有工人的绩效来推断工作任务有多难：如果任务容易，（在任意给定的一组报酬水平上）工人都更有可能市场出更多的产品。问题是，如何才能得到并使用作为激励机制的一个组成部分的信息。我们的分析结果显示，以相对绩效为支付报酬的基础能做到这一点——它是一种相当好的激励机制。我们通过比较为表现最好的工人发奖金与惩罚表现最差的工人这两种方法的优点，进一步研究了最优竞赛的设计问题。

这篇文章同时还有重要的现实意义。在1990年代，在美国及其他一些国家，许多公司的高层管理人员都获得了股票期权。有人认为，股票期权能够为高层管理人员提供良好的激励。但是，事实上，在股票市场繁荣时期，高层管理人员坐在办公室做跟以前一样的事情就获得了报酬；这些报酬基本上没有让他们变得更加努力工作。我们的分析表明，更好的薪酬机制就是以他们的相对绩效为基础支付报酬，例如给在航空业中业绩最好的公司的高层管理人员发奖金，同时要对那些做得最差的企业的高层管理人员罚款。

企业所面临的最重要的一个挑战通常都是对高层管理人员的激励问题。绝大多数首席执行官并不拥有其管理的公司的股票。正如伯利和米恩斯所指出的那样，公司的所有权和控制权是分离的。所有权和控制权分离是第一卷中讨论的更一般的问题——委托代理问题——的一个例子。这个一般性问题通常成为公司治理。正如我在第六卷中所指出的那样，公司治理负面的失败，是俄罗斯和其他东欧与前苏联国家转型失败的原因之一。[7]

在《信贷市场与资本控制》这篇文章中，我最先为此问题的研究奠定了理论基础。我指出，公司对高层管理人员的激励问题，既源于不完美信息

---

7 在捷克共和国，这主要表现为掏空，即管理者或大股东将公司资产转移出去，只留下一个空壳。同时参见 “Quis custodiet ipsos custodies? Corporate Governance Failures in the Transition,” in Governance, Equity and Global Markets, Proceedings from the Annual Bank Conference on Development Economics in Europe, June 1999, Pierre - Alain Muet and J. E. Stiglitz (eds.), Conseil d'Analyse economique, Paris, 2000, pp. 51 - 89，重印于第六卷之中。

问题，又源于良好的管理是公共物品这个事实：也就是说，企业管理得好，企业利润增加之后，所有的股东平等地分享所有收益。因此，任何一个股东都没有足够充分的积极性去监督管理层。当然，如果有一个大股东（比如拥有公司20%以上的股票），那么就有可能减轻搭便车带来的影响（后来的经验研究确证了这个直观的想法）。不过，只有一个控股股东时，就会出现另外一个问题：控股股东将公司利润（或资产）转移到自己名下的动机就会很强烈。此时，保护小股东利益就变得非常关键了。

由于信贷合约的风险较低，所以贷款人不需要承担过高的风险就可以在企业经营管理决策中获得较大的发言权。有了这种地位，它就有积极性去进行监督。同时，银行既有能力去进行监督，也有办法去实现其想法。发放短期贷款的银行，可以拒绝将其贷款展期。由于存在《终止合约的激励效应》所指出的原因，所以公司贷款的银行拒绝为其贷款展期，其他银行通常都不会进行帮助的。于是，公司只能被迫破产。结果就是，虽然从理论上讲，是股权所有者“控制”公司，实际上却是短期信贷资金的提供者实施了有效的控制，而不是股权资金的提供者。

另外一种观点认为，接管能迅速将那些没有最大化公司股价的管理者替换掉。但是，在早期撰写的几篇论文中，我认为接管的作用是很有限的。[8]在《挫败对手》这篇文章中，我将此分析向前推进了一步：我的分析表明，管理者会采取措施进行自我防卫，这就会使解聘他们变得很难。通过将资产投资于价值更加不确定的项目上，来增强在任管理者与外部人之间的信息不对称性，从而使外部人不愿意进行收购。

这篇论文强调了一个一般性的原则：我的许多研究表明，信息不对称会对经济效率和竞争产生负面影响。不过这篇文章强调指出，信息不对称本身可能就是内生的。由于管理者想提高其自身的市场势力，所以许多管理者实际上在那些只能获得较低收益的项目上花了不少钱。

竞争市场理论的一块基石是一价定律：在运转良好的市场上，同一种商品在任何地方都只能以同一个价格销售。一价定律能确保有效配置资源。所有购买者都面临同样的相对价格，就意味着他们会做出相同的相对（边际）

---

8 参见“Some Aspects of the Pure Teory of Corporate Finance: Bankruptcies and Take-Overs,” Bell Journal of Economist, 3（2）, Autumn1972, pp. 458 -482，重印于第三卷之中。以及“Ownership, Control and Efficient Markets: Some Paradoxes in the Theory of Capital Markets,” in Economic Regulation: Essays in Honor of James R. Nelson, Kenneth D. Boyer and William G. Shepherd（eds.）, Michigan State University Press, 1982, pp. 311 -341。

估价，并且会做出同样的权衡取舍，而不管它是消费决策还是生产决策。在完美信息条件下，一价定律显然会成立。如果某处的价格高于另一处的价格（不考虑交通运输成本），那么从价格低的地方买进，到价格高的地方卖出，就可以赚到钱。套利行为会确保每处的价格都是一样的。

但是，许多市场似乎并不像一价定律所描述的那样。答案是显而易见的。市场（或各个小市场）易于受到冲击，而且套利的成本很高。答案依然是一样的：各处的价格差不多会是一样的，而且消除现有的价格差异花费的成本要高于套利得到的收益。

还有人认为，即使有些人的搜寻成本较高，但是还有足够多的人的搜寻成本相对而言比较低，以至于搜寻成本较高的人可以"搭便车"分享由搜寻成本较低的人确定的市场价格秩序。

不过，我在一系列的论文中指出，在信息不完美而且获得信息需要花费成本的条件下，市场实际上是有激励去制造"噪音"。价格分布的出现，不只是因为存在外部冲击，而且还因为有压榨消费者的内在激励。在第五编中的论文（工资和价格分布理论）中，我指出，一般而言市场上都存在价格和工资分布现象；与搜寻成本较低的人相比，搜寻成本较高的人往往要付出较高的平均成本；企业数目较少的市场实际上比那些企业数目较多的市场更有竞争性。分析结果现实，当搜寻需要花费成本时，企业面临的需求曲线就会出现弯折（当价格降低时吸引到的消费者不同于价格升高时失去的消费者）——能导致宏观经济学非常关注的价格刚性出现的弯折。

这一卷最后两编文章主要关注公共政策中的两个重要议题——规制和私有化。在完美信息条件下，规制是很容易的。例如，考虑一个自然垄断的问题。知道了成本和需求曲线，很容易就可以确定出使垄断企业得不到超额利润的价格。但是，监管者通常都不知道垄断企业的成本曲线，而且垄断企业也没有如实披露成本信息的积极性。我很久以前就在论文中指出，从数学上看，最优规制问题等同于最优税收问题。在《信息与规制》一文中，我和戴维·萨平顿（David Sappington）详细阐述了这个一般性原理。

监管显得非常重要的另一个领域是银行业：政府想让银行谨慎行事，以便存款人能取回其存款。同样的，在完美信息条件下，确保银行具有财务自生能力是很容易的；如果没有充分的信息，那就非常困难了。有些人建议要依靠资本充足率。也就是说，只要经过风险调整后的银行资本对其未偿负债（或未偿贷款）的比率满足一定的要求，那么它就能满足兑付要求。充足的资本意味着，如果有些贷款违约了，银行依然能够承担其兑付义务。尤其是

那些倡导金融市场自由化的人认为，银行监管者只需要关注资本充足率即可。在《自由化、银行业道德风险与审慎监管：资本要求管用吗?》一文中，我和凯文·默多克（Kevin Murdoch）以及托马斯·赫曼（Tomas Hellmann）指出，资本充足率只是一个考虑周全的监管体系的一个组成部分而已，并非是唯一的部分；而且提高资本充足率实际上也许会使得银行的行动更加冒险。原因很简单，资本充足率提高之后，银行的“特许权价值”就下降了，（经过风险调整后的）未来利润的预期净现值也会下降。这是因为，银行不得不持有超过其意愿的更多的股权资本。而且，银行之所以有从事谨慎行为的激励，很重要的一个原因就是其享有的银行的特许权价值。

本卷的最后一篇文章考察的是私有化。简单的新自由主义意识形态强调私有化的重要性。它只是简单地假设私营企业比国有企业更有效率。尽管通常都是这样的，但有例子表明国有企业的效率也很高。比如说，韩国的国家经营的钢铁公司的效率远高于美国私营的钢铁公司。加拿大的国家运营的铁路公司的效率与政府经营的铁路公司的效率差异很少。智利的国家经营的铜矿企业据说也与私人铜矿企业一样有效率，但是国有铜矿企业为政府提供了更多的收入。多年前，诺贝尔经济学奖获得者赫伯特·西蒙指出，并不能从理论上说明私营企业一定会比国有企业更有效率——两者都受到我们今天所说的代理问题的困扰。事实上，在本卷前面讨论公司治理问题时，就已经强调一个事实，即不但管理者的利益通常与股东的利益相冲突，而且管理者还会从事许多损害股东利益的事情以增进其自身利益。在《私有化、信息与激励》一文中，我和戴维·萨平顿提了一个简单的问题：在哪些条件下我们才能确保私有化能增进福利？答案并不令人惊讶，它们与保证竞争性市场是有效的市场的条件是一样的——我们已经知道这些条件都是高度抽象的，而且在任何经济体中都得不到满足的，对于发展中国家和转型国家尤其如此。这篇文章进一步支持这么一个研究结论，即并不存在一个一般性的推论：私有化在现实生活中是合意的选择。对私有化效果的衡量还需要逐一考察各个案例。

第二卷中收录的文章只是信息经济学的新范式的广泛应用中的一小部分而已。但是这些论文已经清晰地展现了新范式在帮助我们重新思考旧问题和考察新议题方面的威力。中国在向市场经济转型的过程中，能够在设计其制度、政策和法律体系时借鉴这些新的分析方法，此乃一大幸事。思想观念十分重要。旧的完美信息、完全竞争、完美市场范式对于西方世界形成其制度、政策和法律体系发挥了超乎寻常的影响力，尽管通常都会产

生负面影响。过去，经常用一些七拼八凑的改革措施去应付现实生活中出现的难题。中国现在可以从这些经验中学习，而且新理论既有助于解释早先的经验，又有助于设计促进长期经济增长和建设和谐社会的制度、政策和法律体系。

组织理论

经济体系的结构：等级体系与多极体系*

这篇论文为经济体系和组织的分析提供了一些新的角度和思路。个人的判断往往会出现一些偏差，他们有时会拒绝好的方案而接受坏的方案（或主意）。一个经济体系的结构（例如，体系内部的决策机构是如何组织的，信息的收集者是谁，收集信息的内容是什么，信息的交流者是谁，交流的信息的内容是什么）会影响到体系内个体所犯的错误和这些错误出现的方式。

人们普遍接受这样一个观点：一个经济体系或组织的绩效受其内部结构的影响。在这篇论文中，我们将从一些新的角度来分析经济体系的绩效与经济体系结构某些方面的关系，我们将这类的结构称做经济体系的结构。体系结构（类似于电脑或电气系统的构建结构）描述了体系内部的决策机构是如何组织的，决策权力和能力（ability）是如何在组织内部分配的，信息的收集者是谁，收集信息的内容是什么，信息的交流者是谁，交流的信息的内容是什么。

本文将讨论两种具体的体系机构，即多极体系（Polyarchies）与等级体系（Hierarchies）。我们将多极体系视作拥有几个决策者（决策者之间可能存在竞争关系）的体系，在该体系中决策者可以相互独立地做出决策，决定是否接受某个方案（或主意）。与之相反，在等级体系中决策权是相对集中的，即只有一些人（或者只有一个人）是负责决定是否接受某个方案的，其余的人都

---

* Raaj Kumar Sah, Joseph Stiglitz, The Architecture of Economic Systems: Hierarchies and Polyarchies, *The American Economic Review*, Vol. 76, No. 4 (Sep., 1986), pp. 716 - 727.

为决策者提供支持。这两种体系机构分别代表了市场经济和官僚经济。

我们所关注的组织绩效是指决策的质量。所有的个人都会犯判断的错误：一些被接受的方案实际是应该被拒绝的，而一些被拒绝的方案实际上是应该被接受的。用统计推断中的经典理论做个类比，上述错误分别对应于**第一类错误**和**第二类错误**。

个体的安排方式影响到经济体系所犯错误的性质。例如，在市场经济中，如果一家公司拒绝了一个可盈利的主意（例如，生产一种新的产品），另外一家公司有可能会接受它。相反如果仅由一个部门来做决策，而这个部门否决了这个主意，那么，该主意就完全不会被采用。不过对那些不能获利的主意，情况也是一样的。因此，在多极体系中发生**第二类错误**的概率更大，而在等级体系中发生**第一类错误**的概率更大。

获取和交流信息的成本（导致个体的错误判断）是我们分析时要考虑的重要技术特征，它包括直接成本（时间和资源）以及在信息交流过程中不可避免的信息污染造成的间接成本。交流同决策一样，总不是完美的。没有个体能够完全将其所知道的信息非常完美地传递给别人。

分析时要考虑的另一个重要特征是个体在有限的时间里收集、吸收和处理信息的能力是有限的。因此机构——个体的集合，所能做的可能要比单个个体更多一些（做出的决策也更好一些）。但交流的成本和不完美性也意味着由两个个体组成的组织——例如，每个个体可以在一个月的时间里处理一定量的信息——同拥有在相同时间段里处理两倍信息量能力的单个个体并不一样。

本文的结构如下：在第一部分，我们构造了多极体系和等级体系下决策结构的简单模型。在第二部分，我们假设个体所犯错误的性质和可选方案的集都是外生给定的，并分析了改变这些外生给定的性质会如何影响两个体系的相对绩效。在第三部分，我们讨论在两个体系中对信息的收集和处理过程。尤其是我们将讨论（贝叶斯）甄别原则（screening rules）是如何被决定的以及两个体系的绩效如何同内生决定的个人错误进行比较。在第四部分，我们将简单讨论一些扩展的分析，而在第五部分讨论另一些解释和应用。为了使文章更简洁，大部分结论的证明都省略了，这些证明可以参见我们的工作底稿（1985a）。

一、基本模型

经济体系所面临的问题是要从众多的方案中选出一个实施的方案。每个

方案的净收益为 x[1]，其值可以为正或负。有 N 个可供选择的方案，而方案的密度函数设定为 $g(x)$。

组织中个体的任务是评估（甄别）方案。我们假设个体间唯一可行的交流是根据自己的判断，交流方案是“好”还是“坏”，即该方案应该被接受（或者进入下一轮的评估）还是应该被拒绝（现在，我们可以把筛选者看作是一个黑匣子，当它认为一个方案是好的方案时就会亮闪光）。将给定个体判断某方案为好方案的概率设定为 p，它是方案质量的函数。我们将函数 $p(x)$称做“甄别函数”（Screening function），它可以是任何函数形式，只要对所有的 x 满足 $1 \geqslant p(x) \geqslant 0$ 的条件，并且严格不等式至少对一些 x 成立。

甄别函数有两个特别重要的性质。第一个性质是它的斜率 $p_x(x)$。[2] 我们假设 $p_x(x)$ 为正，即获利越高的方案在甄别中被接受的局部概率也更高。此外，如果 p 和 p^1 代表两种甄别函数，且 $p_x(z) > p_x^1(z)$，那么，我们称前一种甄别函数在 $x = z$ 时是局部更有可判别性的。甄别函数的第二个重要的性质是 $p(x)$ 的大小。如果 $p(z) > p^1(z)$，那么，我们称前一种甄别函数在 $x = z$ 时是局部更松的，而后者是局部更严格的。以线性甄别函数为例，$p(x)$ 可以被表示为：

$$p(x) = p(\mu) + p_x(x - \mu) \tag{1}$$

此处 $\mu = E[x]$ 是初始投资组合（initial portfolio）的均值。很明显，更高的 $p(\mu)$和 p_x 意味着全局更不严格和更好的识别力。

如果甄别是完全的，则体系的结构对产出是没有影响的，因为所有$x > 0$的方案都会被接受，而所有$x < 0$的方案都会被拒绝；也就是说，如果$x > 0$，则 $p(x) = 1$，如果 $x \leqslant 0$，则 $p(x) = 0$。在甄别并非（信息）完全的情况下，经济体系的体系机构决定了甄别方案时的条件，因此，会影响到体系的产出。在下面的简单模型中，我们考虑由两个企业构成的多极体系和由两层机构构成的等级体系。

图 1 和图 2 分别展示了在多极体系和等级体系中的决策过程。在多极体系中，两个企业相互独立地甄别方案。具体而言，我们可认为方案随机地（以 50% 的概率）被两个企业中的一个所选择，如果该方案被其中一家企业

1 该标量值（scalar valuation）包括所有的相关收益和成本。并且，我们假设方案之间的外部性不显著（例如，一个方案的利润不显著地依赖于其他一些方案被实施与否），同时假设可被实施的方案的数目不受限制。

2 下标字母表示对其求偏导的变量。

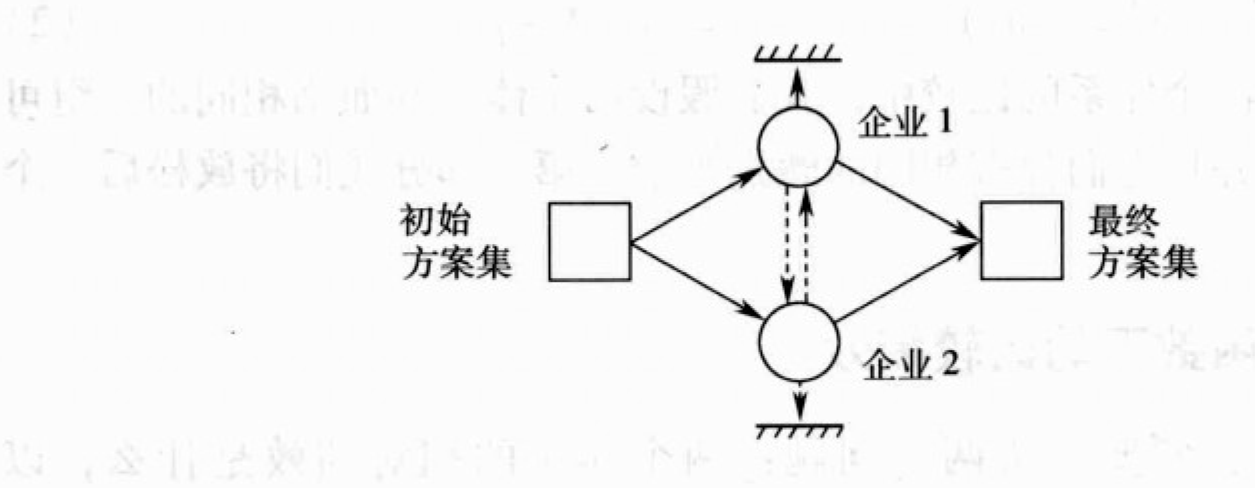


图 1　多极体系

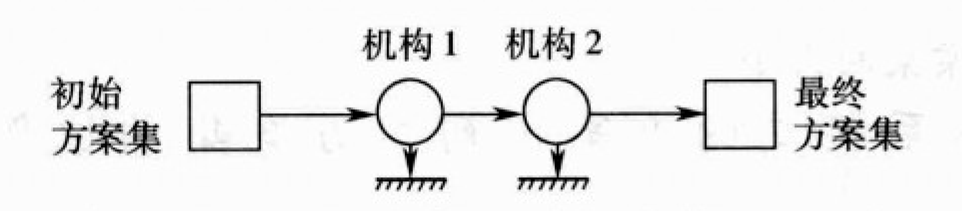


图 2　等级体系

接受了，那它对另一家企业就是不可得的；如果该方案被拒绝了，则其将被另一家企业选择，再一次被接受或被拒绝（但企业不能分辨评估的方案中哪些是已经被评估过的）。任何一家企业都不能对同一个方案甄别两次，因此，同一方案不能在企业之间来回地被甄别。因此，在多极体系中被选择的方案集包括分别被两个企业接受的方案。

与之相反，在等级体系中，所有的方案都先由下层的机构（机构 1）评估，被接受的方案才继续由高层的机构（机构 2）评估，其余的则被拒绝。因此，该体系下被选择的方案是被高层机构接受的方案。我们用继电器电路设计做个类比，在等级体系中甄别是以“串联”的方式安排的，而在多极体系中是以“并联”的方式安排的。

在多极体系中，一个方案被第一家评估它的企业所接受的概率为 $p(x)$，被拒绝的概率为 $1-p(x)$，而其再被第二家企业接受的概率也为 $p(x)$。因此，被接受的总概率为 $p(x)+(1-p(x))p(x)=p(x)(2-p(x))$。与之相似，在等级体系中，一个方案被低层机构接受的概率也为 $p(x)$，而该方案被高层机构接受的概率也为 $p(x)$，因此，该方案最后被接受的概率为 $p^2(x)$。

方案 x 被体系 s 所接受的概率可以表示为 $f^s(x)$，此处的上标 s 为 P 或 H，分别表示多极体系和等级体系。如果个人的决策在甄别和方案上都相互独立的，那么

$$f^H \equiv (p^H)^2; \qquad f^P \equiv p^P(2 - p^P) \tag{2}$$

在下面的对两个体系的比较中，我们假设两个体系都面临相同的一组可供选择的方案，并且它们都有相同的甄别函数。第三部分我们将放松后一个假设。

二、相同甄别函数下的比较绩效

在本部分，我们要分析两个问题：两个体系的相对绩效是什么，以及可选方案集（project portfolio）的性质和甄别函数是如何影响相对绩效的？

1. 最终方案集的大小

被两个体系接受的方案占初始方案集的比例 n^s，正好为 $\int f^s(x)g(x)dx \equiv E[f^s]$。两个体系接受比例的差可用 Δn 来表示，我们发现[3]

$$\Delta n = n^P - n^H > 0 \tag{3}$$

因为从式（2）可得 $f^P - f^H = 2p(x)\{1 - p(x)\} \geq 0$ 对所有 x 都成立，而对一些 x 其值严格为正。

定理1：相对等级体系而言，多极体系接受的可选方案的比例更大。

该结论背后的逻辑是非常直观的。考虑一种假设的情形：在多极体系中，不存在第二家公司，在等级体系中也不存在高层的机构，则在两个体系中，被接受方案的比重将是一样的，也就是 $E[p(x)]$。由于第二家企业至少会接受一些方案，而高层的机构至少会拒绝一些方案，因此，在多极体系中被接受方案的比重肯定要高于在等级体系中被接受方案的比重。同样直观的是该结论不仅对好方案成立，对坏方案也同样成立。进而言之，该结论并不依赖于如何定义好和坏方案，只要在甄别中存在至少接受一些好方案和一些坏方案的概率。紧接着我们可以得到的结论为：相对于等级体系而言，无论如何定义好方案和坏方案，多极体系接受的好方案和坏方案的比例更大。因此，在等级体系中发生**第一类错误**的概率相对更高，而在多极体系中发生**第二类错误**的概率相对更高。

以上结论表明在一些情况下，多极体系的表现会比等级体系好（当避

3　此处同别处一样，可选方案的数目 N 无关紧要，因此我们将其省略了。已用标准化的可选方案的数目来代表一个体系绩效的变量。

免**第一类错误**更重要时），而在另外一些情况下，等级体系的表现会比多极体系好（当避免**第二类错误**更重要时）。

在分析初始方案集的性质对最终方案集大小的影响时，从式（2）中我们注意到$f^s(x)$对x是增函数。此外，如果甄别方程是线性的，则f^P是x的凹函数，f^H是x的凸函数。因此由统计占优的标准性质（假设方案分布的端点是固定的）我们得到下面的结论：

定理2：初始方案集的一阶随机占优的减低，导致在两个体系中初始方案集被选择的比重的降低。当甄别函数为线性时，初始方案集的均值保留展型（mean-preserving spread）将导致在多极体系中可选择方案的比重减小，而在等级体系中被选择方案的比重增加。

图3展现了上述结论。如图3所示，因为筛选方程是线性的，则f^P是x的凹函数，f^H是x的凸函数。n^s是x轴之上以f^s和g的乘积为边界的区域。自然地，对应于f^P的该区域比对应于f^H的区域大；并且如果密度权重从较低的x值移向较高的x值，多极体系和等级体系的该区域都会增大。如果由于均值保留展型，且密度权重从中值移向两边，则对应于n^s的区域在多极体系中减小，在等级体系中增大。

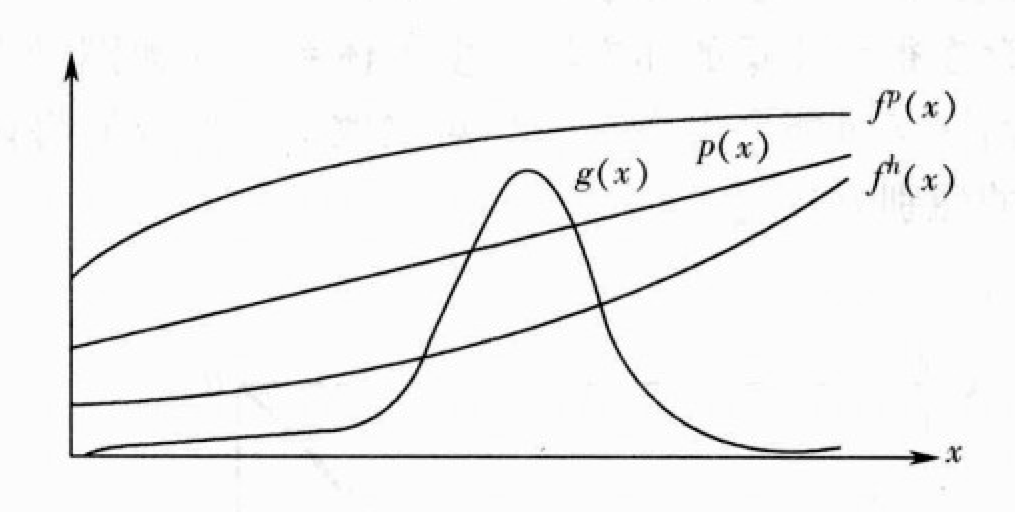


图3　多极体系和等级体系被接受的概率

进行一些简单的计算可以使我们确知线性甄别方程中的两个参数的变化是如何影响n^s的。我们发现如下定理：

定理3：当甄别方程为线性的时候，甄别越不严格（$p(\mu)$更高），越会导致两个体系中被接受方案的比例增大。而甄别的识别力越高（p_x更高），越会降低多极体系中被接受方案的比例，而增加等级体系中被接受方案的比例。

2. 两个体系的利润

两种类型的方案。当在被评估的方案集只由两种类型的方案构成时，我

们可以完整地描述出在哪些情况下多极体系能够获得更高的（期望）产出，哪些情况下等级体系能够获得更高的（期望）产出。初始方案集可由三个变量来描述：好方案的回报 $z_1>0$；坏方案的回报 $-z_2<0$；以及好方案所占的比重 α。用好方案被接受的概率 $p_1 = p(z_1)$ 和坏方案被接受的概率 $p_2 = p(-z_2)$ 来描述甄别方程。如果 $Y^s = E[xf^s]$ 表示产出，$\Delta Y = Y^P - Y^H$ 表示两种体系产出之间的差别，那么

$$\Delta Y = 2z_2(1-\alpha) \times [ap_1(1-p_1) - p_2(1-p_2)] \tag{4}$$

此处 $a = z_1\alpha/z_2\ (1-\alpha)$ 表示初始方案集的整体质量。

在此模型中初始方案集的质量的提高表现为 a 的值更高（也就是，α 的值更大或者z_1/z_2的值更大）。由式（4）可知初始方案集的质量的变差意味着多极体系相对于等级体系的相对绩效变差。这点是非常直观的，因为等级体系的相对优势在拒绝坏方案，而多极体系的相对优势在接受好方案。如果初始方案集质量的变差，则前一种优势的重要性递增而等级体系的相对绩效会提高。然而我们必须注意除了别的一些因素，一个方案在甄别过程中被接受或被拒绝的概率（也就是，项目被接受或被拒绝所遵循的规则），还可能会被可供选择的方案集的组合状况所影响。例如，有人可能会猜想如果存在很大比重的坏方案，多极体系中的甄别过程可能会变得相对更严格，而这又可能会提高它的相对绩效。后面我们将讨论具有该种性质的内生性的甄别方程。

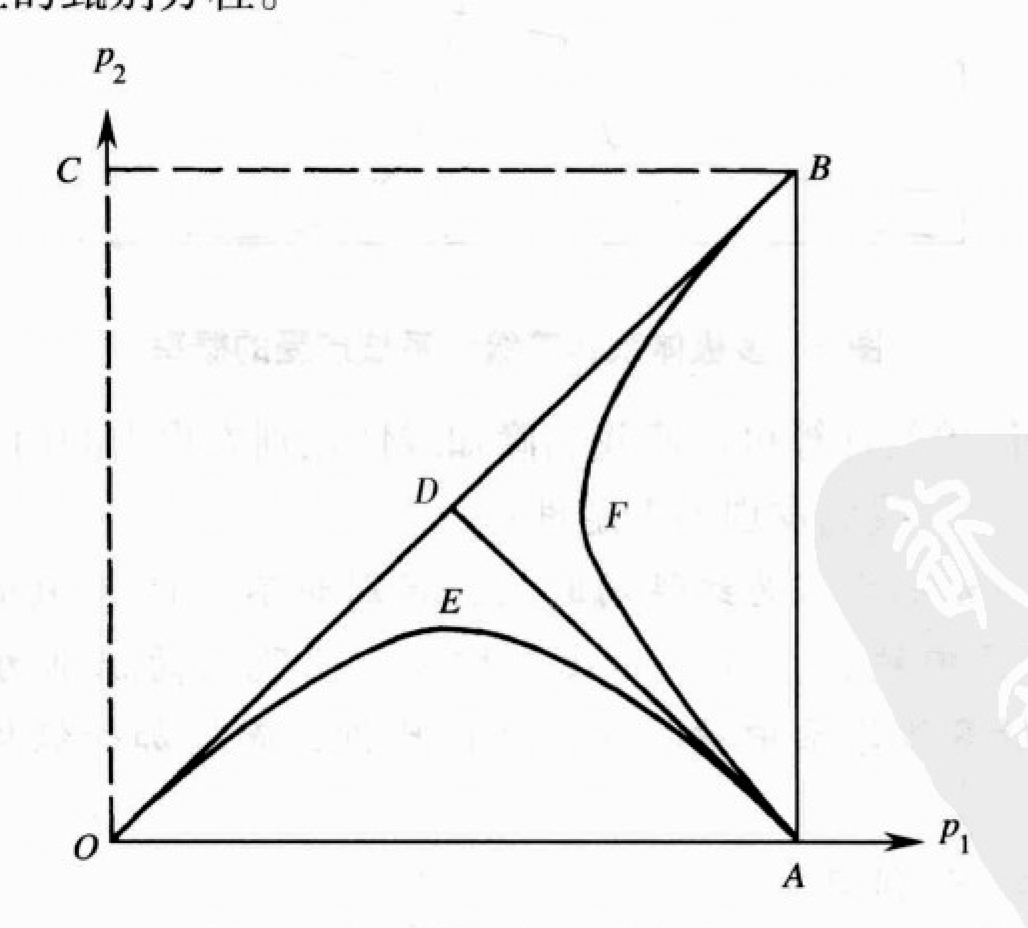


图 4　多极体系和等级体系比较

上面的表达式允许我们把参数空间划分为两个区域：一个是多极体系比等级体系的产出要高的区域，一个是相反情况的区域。

图4展示了一些结论。我们只关心45°线以下的区域，因为甄别应具有一定的识别力，即$p_1 > p_2$。

首先考虑当初始方案集的质量一般好的情况，即$a = 1$，例如，初始方案集的好方案和坏方案的个数一样（$\alpha = 1/2$），并且从两类方案中获得的收益和损失是对称的（$z_1 = z_2$）。在这种情况下，如果满足下列条件，多极体系的产出更高：

$$1 - p_1 > p_2 \tag{5}$$

因此在图4中，多极体系在区域*ODA*中的表现比等级体系更好，而在区域*ADB*中则相反。

对该结论的解释非常简单。由前文可知（$1 - p_1$）是甄别时犯第一类错误的概率，即拒绝好方案的概率；而p_2是筛选时犯的第二类错误的概率，即接受坏方案的概率。现在，如果在甄别时拒绝好方案的概率大于接受坏方案的概率，也就是如果式（5）成立，那么多极体系（该体系给了被拒绝的方案第二次的机会）的绩效会更好。

如果初始方案集的质量较差，也就是$a < 1$，那么从式（4）可知参数空间被位于*ODA*区域内的双曲线*OEA*分割了。在*OEA*区域内，多极体系获得的利润更高，而在该区域外则相反。*OEA*区域随着初始方案集质量变差而变小，当$a \approx 0$时，它和*OA*线重合。在相反的情况下，即初始方案集的质量较好（例如$a > 1$），也有对应的结论——在*AFB*区域外，多极体系有更高的利润，而在该区域内，情况则相反。

我们可以用另一种直观的方法来解释这些结论。假设我们使每个方案都经历两次的甄别。如果两次甄别都认为该方案是坏的，该方案将会被拒绝，而如果两次甄别都认为该方案是好的，则该方案将被接受。当对该方案的看法不同时就会出现权衡（tradeoff）。有争议的方案是否应该被接受依赖于从该方案中所能获得的利润。好方案引起争论的概率为$2p_1(1 - p_1)$，而坏方案引起争论的概率为$2p_2(1 - p_2)$，因此，有争议的方案能够带来的期望利润和式（4）的一样。如果式（4）的值是正的，意味着有争议的方案应该被接受，而这正是多极体系所能保证的。同样地，如果式（4）的值是负的，则意味着有争议的方案应该被拒绝，而这正是等级体系所能保证的。

一般性的方案集。在结束本小节之前，我们简略地分析一下由连续质量

的方案构成的初始方案集。由式（2）可知 $f^{P}-f^{H}=2p(1-p)$，而 $\Delta Y=2E[x\psi]$，此处的 $\psi=p(1-p)$。为了便于分析初始方案集的变化的影响，我们用以下的表达式：

$$\psi_{x}=(1-2p)p_{x};\psi_{xx}=-2p_{x}^{2}+(1-2p)p_{xx} \tag{6}$$

由于 ψ 可以是 x 的凹函数或者凸函数，总体而言，初始方案集的均值保留展型变化的影响是不明显的。如果 x 的值域较小，那么用式（6）我们可以得到：

$$\begin{aligned} E[x\psi] &\approx \psi(\mu)\mu+\psi_{x}(\mu)E[x(x-\mu)] \\ &= p(\mu)(1-p(\mu))\mu+(1-2p(\mu))\times p_{x}(\mu)E[(x-\mu)^{2}] \end{aligned} \tag{7}$$

因此，如果 $\mu\geqslant 0$ 并且 $p(\mu)<\frac{1}{2}$，那么，多极体系的产出就比等级体系的大；进而，无论 μ 的值为多少，方案集的质量方差的增加会提高多极体系的相对绩效。

当甄别方程为线性时，我们很容易得到另外一些比较静态的结论。例如，如果初始方案集所包含的方案以 0 为均值对称分布，那么，多极体系比等级体系的表现更好还是更差简单地依赖于 $p(\mu)$ 的值是小于还是大于 0.5，也即甄别是否严格。另一方面，如果初始方案集有更高的均值或者更大的负偏态，将提高多极体系的相对绩效。

三、内生性甄别机制

以上模型中的个体决策者可以被视作是服从贝叶斯判决规则的，每个人都接收到关于方案质量的（不完美的）二元信号。更一般的情形是个体观察到多个需要他们理解的信号；换言之，他们需要决定在何种情形下会建议接受某个方案。例如，假设方案评估者观察到：

$$y=x+\theta \tag{8}$$

方案评估者用保留水平（reservation level）的方法来筛选：如果一个方案被观察到的利润高于保留水平 R，就接受，否则就被拒绝。[4] 假设 θ 的分布独立于 x，其分布函数为 $M(\theta)$，密度函数为 $m(\theta)$。那么，甄别函数可以写为：

$$p(x,R)\equiv \mathrm{Prob}[y\geqslant R]=1-M(R-x) \tag{9}$$

由以上的表达式可得 $p_{x}\geqslant 0$，而 $p_{R}\leqslant 0$：一个方案在甄别中被接受的概率是该方案质量的增函数，是保留水平的减函数。递增的 R 会增加好方案被拒

4 最优政策可以单独用保留水平来描述，只要残差项的性质满足一些适度正则条件。

绝的概率（**第一类错误**）而降低坏项目被接受的概率（**第二类错误**）。人们通过选择保留水平 R 平衡这两种错误。

在多极体系中，用上标 i 和 j 来代表两家企业。对企业 i，R^i 是保留水平，$p^i \equiv p(x, R^i)$ 为甄别函数，Y^{iP} 为产出，那么：

$$Y^{iP} = E[xp^i(2 - p^j)]/2; Y^P = Y^{1P} + Y^{2P} \tag{10}$$

接下来我们比较两种体系下的保留水平，并用这些结论来比较两体系的绩效。在等级体系中的保留水平 R^H，为最大化 $Y^H = E[xp^2]$，它满足下式：

$$Y_R^H = 2E[xpp_R] = 0 \tag{11}$$

为了强调多极体系中两家企业的独立性（并有潜在的竞争性），我们假设它们的保留水平的制定是没有经过相互协调的。我们关注于对称的纳什最优，从式（10）可知，对应的保留水平 R^P 可由下式得到：

$$E[x(2 - p)p_R] = 0 \tag{12}$$

我们还发现在“协调过的多极体系”（在这种情况下，企业的保留水平是根据两家企业总产出最大化的原则制定的）中，作为基准的保留水平 R^C 可以通过使下式等于零得到：

$$Y_R^P = 2E[x(1 - p)p_R] \tag{13}$$

由式（11）到式（13），我们可以得到如下定理：[5]

定理4：$R^C > R^P > R^H$。在多极体系中的甄别比等级体系中的更为保守，但比协调过的多极体系中的甄别的保守程度更低。

对这个结论的解释很直观。在等级体系中，低层机构知道它的决策还会被高层结构再核查，而上层机构知道所有它接到的方案都已经被低层机构核查过了；在多极体系中，每家企业其决策不会再被核查了，对这些企业更糟的是，它知道评估的方案中包括许多已经被别的企业核查过并拒绝了的方案。在市场经济中的企业，实施方案时要求高的“期望”回报就反映了这种保守性。例如，企业往往会设定这样的决策原则——只有期望回报率超过20%的方案才会被采用，但实际的平均回报率是相当小的。企业知道要达到

5　为证明该定理，我们定义 $c(x) = (1 - p(x))/p(x)$，并发现对任意常数 $c(0)$，$c(0)Y^H$ 的拐点和全局最大值和 Y^H 的是相同的，而且 $d(c(0)Y^H - Y^P)/dR = 2Expp_R[c(0) - c(x)] < 0$，因为，$c_x(x) < 0$，$p_R < 0$，假设是相反的结论，即 $R^H > R^C$。那么从 $Y_R^P > c(0)Y_R^H$ 可得 $Y^P(R^H) - Y^P(R^C) > c(0)\{Y^H(R^H) - Y^H(R^C)\}$。不等式的左边为负值，而右边为正值，该矛盾说明了 $R^C > R^H$。非协调的多极体系可以认为是以最大化 $Ex[2p - .5p^2]$。证明结论 $R^C > R^P > R^H$ 的方法和证明 $R^C > R^H$ 的方法类似。

要求的回报率，它们就必须设定高的保留水平。正如我们事前所猜想的那样，我们的分析也说明了，与协调过保留水平的企业相比，在非协调的多极体系中企业不会考虑它们对其他企业造成的负的外部性（例如每家企业都使其他企业所面临的可选方案的质量下降了）。

保留水平的比较静态分析。由式（11）到式（13）可以立即得到下面的定理：

定理5：坏方案的相对比重的显著增加会使多极体系和等级体系的保留水平都增加。即如果当 $x \geqslant 0$ 时，$\partial g(x,\beta)/\partial\beta \leqslant 0$，则 $dR^s/d\beta > 0$，对 $s = P, H, C$ 成立。

因此，可供选择的方案集的质量变差会导致甄别更严格（更低的接受概率）。

决定保留水平的另一个重要因素是选择方案时所依赖信息的质量。直观地，我们可以预期信息的质量变差会导致更高的保留水平。简而言之，使方案评估者对项目的观察方程中的残差项和 x 相关。考虑 θ 服从另一种分布的情况，当 $x < 0$ 时，该分布一阶随机占优于原来的分布。因为在新的分布下，我们在任何的保留水平下接受坏的方案的可能性都增加了，因此，我们认为新的信息比原信息具有更多的干扰性。由此我们得到：

定理6：如果甄别是严格的，更不确定的信息导致等级体系的保留水平增加。[6]

绩效比较。由于在不同体系中保留水平的设定是不同的，个体所犯的第一类错误和第二类错误也不同，这使得体系绩效的比较比在前面的比较要困难些。主要分析只有两种方案的特殊情况，我们得到了三组结论。

（a）假设当两个体系都使用等级体系的（最优）保留水平，则多极体系比等级体系的产出要大。并且很明显地，用自身保留水平协调的多极体系的表现会更好。此外，如果 Y^P 是 R 的局部凹函数，[7] 那么由定理4可得，使用自身保留水平的非协调的多极体系的表现也比等级体系的好。

因此，当 $a=1$，并且等级体系的保留水平使得甄别概率落在图4中的

---

6 此处对不确定的增加的定义和统计决策文献中的标准定义不同。在上述的条件下，对所有 $x<0$ 的值，$p(x,R)$ 是 R 的增函数，而 $p_R(x,R)$ 在 R 最优值处递减，只要残差项的密度函数对较大的 θ 值是递增的。该结论使 Y^H 的每个局部最大值往右移。为了确保最大值减小了，我们需要假设凹度，或对干扰项分布的变化的条件做一些更强的假设。

7 即在 $R^H < R < R^C$ 的范围内，$Y^P_{RR} < 0$。

ODA 区域中，**那么多极体系的产出更大**。同样的道理，当多极体系的保留水平使得甄别概率落在 *ADB* 区域内，那么等级体系的产出更大。当 $a \neq 1$ 时也可以做类似的解释。值得注意的是如果等级体系的甄别概率在 *ADB* 区域内或者多极体系的甄别概率在 *ODA* 区域内，那么，我们就得不到关于相对绩效的判断（使用这种方法）。

在何种情况下等级体系的甄别概率是使多极体系的绩效好于等级体系时的概率？首先，考虑 a 略微小于 1 的情况。那么，当我们观察到完全不具有信息的信号时，$p_1 = p_2 = 0$。由上面的结论和连续性可知，**如果关于方案质量的信号足够的差，那么多极体系比等级体系的绩效好**。这个结论稍微让人有点惊讶：有人可能认为当信息很差时，等级体系下的第二次甄别会更有价值。但在保留水平调整很多，决定性甄别很严的条件下，多极体系中的第二次机会比等级体系中第二次评估更重要。

其次，从定理 5 可知方案集质量的提高（例如，a 的增加）会导致更高的接受概率。另一方面，我们前面的分析证明了更高的 a 扩展了使多极体系优于等级体系的 p_1 和 p_2 的值域。因此，我们确定了多极体系绩效优于等级体系的条件。等级体系内的最优，式（11），在这种情形下（当初始方案集只由两种方案构成）被重新表述为：

$$ap_1 p_{1R} = p_2 p_{2R} \tag{14}$$

再定义 $k \equiv [m(R^H - z_1)/M(R^H - z_1)]/[m(R^H + z_2)/M(R^H + z_2)]$

由式（4）可以得到以下结论：

定理 7：如果 $k \leqslant 1$，多极体系的绩效要好于等级体系的绩效。

接下来必须考虑 $m(\theta)$ 为单峰分布和 a 的值较大的情况，因此需要保留水平足够低，使得 $R + z_2$ 和 $R - z_1$ 都低于众数。那么倘若在 $R - z_1 < \theta < R + z_2$ 时，$m(\theta)/M(\theta)$ 是 θ 的增函数（也就是，$m_\theta > m^2/M$）。通过证明在何种情况下，等级体系的绩效在多极体系的最优甄别概率处优于多极体系，也可以得到类似的结论。

（b）我们可以推导出另一些较弱的充分条件，在这些条件下，考虑到多极体系下的甄别更严格的事实，任一体系的绩效都会更好。让一个多极体系选择一个保留水平，在该保留水平下多极体系实施的方案的（预期）数目和一个等级体系所选择实施的最优数目是相同的。该保留水平一定不是多极体系的最优保留水平，但如果我们能够证明在这个非最优的保留水平下，多极体系的绩效要优于等级体系，那么，一个协调过的多极体系在选择其最优保留水平后，绩效一定会更优。

如果 R^T 表示多极体系使用的保留水平，在该保留水平下，多极体系和使用 R^H 作为保留水平的等级体系选择的方案数目一样多，那么 $R^T > R^H$。为简单起见，我们同样定义 $m^P(x) \equiv m(R^T - x)$ 和 $m^H(x) \equiv m(R^H - x)$，以及 M^H 和 M^P 的定义类似。如果该多极体系的甄别函数更具有识别力，也即在一定的范围内，[8] $f_x^P > f_x^H$（等价于如果 $m^P M^P > m^H(1 - M^H)$），那么，该多极体系的绩效优于等级体系。如果甄别十分严格（M^H 的值接近 1），只要 m^P 和 m^H 的差值不太大——如果 m_θ 的值不太大，该条件就成立——该结论就成立。进而言之，当甄别非常严格时，$R^P \approx R^C$（例如，外部效应可以忽略不计）。因此，（在上述条件下）协调的多极体系和非协调的多极体系的绩效都要优于等级体系。

如果 R^T 并没有比 R^H 大很多，就需要建立更弱一点的条件。这种情况下的充分条件是甄别过程严格程度一般，即甄别方案的接受概率小于 0.5。[9]

（c）我们已经详细地讨论了当残差项的分布为均值为 0 的均匀分布，方案质量对称（$z_1 = z_2$）时的情况。在这种情况下，**如果初始方案集中好方案的比重小于（大于）0.5，则多极体系的利润比等级体系的利润高（低）。**很明显，如果我们假设方案集中亏损的方案明显地要比盈利的方案多，那么，目前的范例说明多极体系是更优的制度安排。

评注：

（i）有限交流的假设在我们的分析中扮演了很重要的角色。我们假设在多极体系中的企业完全不能（或不）交流，而在等级体系中的机构之间也只交流二元信号（它们认为一个方案是好的还是坏的）；它们之间不能交流它们所了解的关于方案特征的真实信息。我们相信在不同的条件下信息分享的程度是不一样的，但极少会是信息交流完全的，而我们构造的模型即用于分析这点所导致的结果。

（ii）经济体系自身的结构揭露了一些关于它的构成要素的信息，这些

---

8 某个方案被多极体系接受的概率为 $f^P(R^T, x) = 1 - M^2(R^T - x)$，此处分别对应于好方案和坏方案 $x = z_1$ 和 $-z_2$。在等级体系中相应的概率是 $f^H(R^H, x) = \{1 - M(R^H - x)\}^2$。在两个体系中，相同数目的方案被接受，那么，如果 $f^P(R^T, z_1) > f^H(R^H, z_1)$ 或者，$f^P(R^T, -z_2) < (f^H, -z_2)$，多极体系的绩效高于等级体系。等价的条件是 $f^P(R^T, z_1) - f^P(R^T, -z_2) > f^H(R^H, z_1) - f^H(R^H, -z_2)$，当 $f_x^P > f_x^H$ 在一定范围内成立时，前式成立。

9 这是因为如果 $R^T \approx R^H$，那么，只要 M 大于 1/2，在 $R^H - z_1$ 到 $R^T + z_2$ 的范围内，要求的不等式［参见式（8）］$f^P(R^T, z_1) - f^P(R^T, -z_2) > f^H(R^H, z_1) - f^H(R^H, -z_2)$ 成立。

信息可以被用来制定决策规则。例如，在我们分析等级体系时，体系内的企业之间不分享任何的信息，每家企业都知道它所接到的方案中有一些是被其他企业所拒绝的，因此，每家企业面临的方案集都不是初始方案集的翻版，而是被其他企业更改过的。在制定最优保留水平的时候就要部分用到这一潜在信息。

四、扩展

我们的模型的基本要素包括甄别函数、可选方案的分布、体系的结构和体系自身的决策规则。每个要素都可以将其一般化（在我们引用的文章中已经扩展了部分）。在接下来的这个部分中，我们在给定已有信息的条件下将甄别函数内生化。如果把更多的资源分配到信息的获得上，可以获得更多的有信息的信号。那么，等级体系中的每一级（多极体系中的每家企业）在信息获得上的花费必须是内生决定的。一般而言，等级体系中每一级在信息获得上的花费是不一样的：低层机构的高质量甄别提高了被高层机构评估的方案集的质量，但由于低层机构要评估大量的方案，花费也较高。当在等级体系中分配具有不同能力的个体时也会出现类似的情况（在这篇论文中假设所有的个体的能力都是一样的）。体系的结构可能也会影响到可选方案的组合，因为不同类型方案被接受的可能性也可能影响到研究（评估）动机。

我们研究过的另一种“体系机构”是委员会，即由个体组成大小不同的团体，使用特定的决策规则（例如，多数投票制）来批准方案（这篇论文讨论的决策单位可以被看作是只有一个成员的委员会的极端例子）。反过来，这些极端的结构可以被看作是拥有等级体系、多极体系和委员会的混合结构的复杂组织和经济体的基本构成要素。

我们也可以来讨论一下不同的决策规则所带来的结果。我们遵从了一个自然的假设：一个方案只有在被组织批准了才可以实施；在等级体系中要被两级机构批准；在多极体系中要被至少一家企业批准。当然我们也可以想象一些完全不同的决策方式：例如，在等级体系中所有的方案都通过了，除了那些被两级机构都否决了的方案；而在多极体系中，一个方案只要被一家企业否决了就不被接受。

我们将那些根据否决原则进行决策的组织称为否决型等级体系和多极体系，它们和我们之前分析的可称做接受型等级体系和多极体系的组织形成了对比。很容易证明，在不存在协调成本的情况下，接受型的多极体系（等级体系）和否决型的等级体系（多极体系）是等价的。

然而，值得注意的是不同体系的协调要求可能有显著的不同，这依赖于体系的结构和决策规程的性质。例如，接受型多极体系不要求构成单位之间进行任何的信息协调，一家企业不需要通知其他企业它接受或拒绝了某个方案。在类似市场的体系中，这是企业之间相互独立的一个重要方面，这点我们以前就已经强调过了。与之相对应，在否决型各级体系中，每个决策单位都可以否决一个方案的实施，因此，每个单位必须通知其他单位它拒绝了哪个方案。如果信息协调的成本很高（在信息传递过程中的干扰很大），那么，接受型的多极体系可能要比其他的组织形式更有优势一些。

五、应用和结论

我们分析的重点在于不同的形成决策的方法，这些方法也许可以被任何一个组织所采用。对组织的选择可以是明确的或是暗含的，而这些比较静态定理揭示了当我们观察到某种形式的组织时，它所处的“客观”环境。这样的比较分析不仅同企业决策有关，还同公共部门的决策有关：例如，持续的关于军队组织机构的争论（是三个相对对立的陆海空部队还是一个统一的陆海空三军）。类似地，不同的政治结构（拥有不同的核查和制衡体系）可以被看作是不同的体系机构，都是用来平衡人类可能犯的不同错误所导致的后果的。

在文章的开始部分即提到了我们的方法的一个很重要的应用，即用来比较组织经济体系的不同方法，多极结构表现了市场经济一些核心要素，而等级结构表现了集权经济体的核心要素。决策的组织形式（以及对应的错误、成本和后果）在这个研究领域之前的大多数重要研究中都没有得到重视。对经济体系的选择，显然要比诸如对不同的一劳永逸的社会资源的分配方法的比较要危险得多，认为市场社会主义与市场经济可以具有等效性的兰格—勒纳—泰勒（Lange-Lerner-Taylor）观点，即强调该点。

我们对于不同体系比较的很多方面都还没有在此进行充分的讨论，但我们相信它们是可以在我们的分析框架内被分析的，这也是我们以后要进一步做的工作：例如，认为在分权式的经济体系中常常能有更好的激励机制的观点（Barry Nalebuff and Stiglitz，1983），认为集权经济体系能更好地处理外部性问题的观点（我们在分析中已经明确排除了），认为分权体系的绩效对关键决策者的品质不够敏感的观点，认为自然选择机制在分权经济体中起到更有效的作用的观点，以及认为分权结构为创新提供更大的激励的观点。

但是，我们的分析围绕关于多极体系和等级体系的争论，讨论了其他的

几个方面：多极体系的优势在于好方案有更多的机会被接受，而其劣势在于不能对不合格的决策提供足够的核查。而对等级体系的批评又在于它的核查成本太高了，既存在由于更多的评估导致的直接成本，也存在由于在处理过程中为了确保坏方案不被实施而导致好方案也被拒绝的间接成本。对多极体系的支持性观点进一步指出它在交流的经济节约方面的价值。这些观点都具有真理的种子。在这篇论文中，我们提供了一个分析框架，在这个框架内，人们可以评估在哪种条件下，某一种观点比另一种观点更有道理。

参考文献

Lange, Oskar and Taylor, Fred M., *On the Economic Theory of Socialism*, New York: McGraw-Hill, 1964.

Nalebuff, Barry and Stiglitz, Joseph E., "Information, Competition and Markets," *American Economic Review Proceedings*, May 1983, 73, 278 – 83.

Sah, Raaj Kumar and Stiglitz, Joseph E., (1985a) "The Architecture of Economic Systems: Hierarchies and Polyarchies," Economic Growth Center Discussion paper No. 476, Yale University, 1985.

______ and ______, (1985b) "Perpetuation, Self-Reproduction, and Evolution of Organizations: The Selection and Performance of Managers," paper presented at the 1985 World Congress of the Econometric Society, Cambridge, August 1985.

______ and ______, (1985c) "Economics of Committees," Economic Growth Center Discussion Paper No. 486, Yale University, 1985.

人类易错与经济组织*

与其他任何学说相比，关于社会组织的好学说对人类社会的影响最为深刻。具体来说，人们认为某种生产和交换模式好于其他模式的观点激发了一系列社会经济试验，导致了现代资本主义社会和社会主义社会的产生，并且意义深远。然而令人费解的是，人们对经济组织中心学说（虽然一直是意识形态辩论的源泉）的科学研究却十分有限。主要的假设，也就是所谓的兰格—勒纳—泰勒定理认为，竞争的资本主义经济与分散的社会主义经济在运用价格体系方面是等效的，它认为生产方法的所有权问题并不处在比较经济制度的核心位置。然而，定理是建立在资本主义和市场社会主义的模型之上，模型中两种制度最重要的区别却被忽略了。

本文阐述了一个研究项目，试图描述经济组织替代模式的一些根本性差异。我们认为，理解这些差异的核心是理解组织决策的差异；以及理解谁搜集什么信息，如何交流，与谁交流和如何做决策的问题，上述两方面都涉及采取什么行动以及谁来决策的问题。这一观点与决策不起作用的传统经济框架相悖：例如，经理人照本宣科地简单地查询在给定的要素价格下生产的适宜技术。在传统的框架下，无论是搜集信息还是传递信息或者做决策，人们从不犯错。此外，事实上，这些行为没有成本。相对而言，我们的观点是"人必犯错"，不同的组织体系存在差异，不仅在于个体犯何种错误，也在于系统是如何"加总"错误的。其结果是，不同的组织在所犯的错误上存在系统性差异，因此，它们在总体经济运行上也存在差异。不同的组织在与成本相关的信息搜集、信息交流和处理，以及决策上也存在差异。事实上，如下所述，我们可以通过合适的方式安排足够多的决策人来实现完美决策，无论决策人有多大的失误，只要他们的决策不是完全随机的（或者更糟）。成本阻碍完美决策。

我们把信息搜集、交流和决策的具体结构作为组织的架构

---

* Raaj Kumar Sah, Joseph E. Stiglitz, "Human Fallibility and Economic Organization", *The American Economic Review*, Vol. 75, No. 2, Papers and Proceedings of the Ninety-Seventh Annual Meeting of the American Economic Association (May, 1985), pp. 292 - 297.

(architecture)。我们研究项目的目的是建立系统化的经济组织模型，来验证替代的组织架构的重要性。

运用这些模型，我们不仅可以对特定组织形式的运行进行比较，我们也可以发问，给定特定的目标和情形集合，最优的结构是什么（在一类结构中)。例如，等级体系的最优级别数是多少。我们建立了实际和标准化两种理论。因此，我们的方法允许我们去评估许多传统上主张的有关替代系统的有效性。

一、基本模型

我们用最简单、可行的公式来阐述讨论问题的基本原理。考虑一个组织面临这样一种问题，它要从大量可利用的两种类型项目中选择：预期收益 x_1 的好项目和预期收益 $-x_2$ 的坏项目。α 比例的项目为好项目。在信息充分条件下，所有好的项目将被采纳，没有坏的项目被采纳。决策人根据自己掌握的信息来判断项目是好是坏。我们假设，一个决策人不能完全与他人交流已有信息。这里我们描述一个“有限”交流的极端形式，也就是说，在任何组织中的个体虽然有很多信息，但他们仅互相交流对项目好坏判断的信息。

所有的决策是不完美的；我们假设决策人判断一个好项目是好的概率为 $p_1<1$，他判断一个坏的项目是好的概率是 $p_2>0$；存在一些过滤的事实反映在 $p_2<p_1$。

我们考虑两个不同的组织，每个由两个个体组成。在多极体系中，每一个个体有权接受一个项目，以及由一人拒绝的项目可由他人来评估。因此，好项目被接受的概率是 p_1（在第一次评估中接受的概率）加$(1-p_1)$（在第一次评估中拒绝的概率）乘 p_1，在第二次评估中被接受的概率。因此好项目被接受的概率是$f_1^P=p_1(2-p_1)$。同样地，一个坏的项目被接受的概率是$f_2^P=p_2(2-p_2)$。组织的产出（期望值）是 $Y^P=\alpha x_1 f_1^P-(1-\alpha)x_2 f_2^P$。

比较而言，在等级体系下，项目被采纳，它必须被上下两级批准，好项目被批准的概率是$f_1^H=p_1^2$，坏项目被批准的概率是$f_2^H=p_2^2$。等级体系组织的产出是 $Y^H=\alpha x_1 f_1^H-(1-\alpha)x_2 f_2^H$。

我们随即得出两个结论：多极体系接受较坏的项目（$f_2^P>f_2^H$）；等级组织拒绝更多好的项目（$f_1^H<f_1^P$）。两种体系犯不同错误的事实意味着存在一个比另一个运行更好的情形。我们可以通过比较两种体系的净产出来探知那些情形：

当 $\alpha x_1 p_1(1-p_1) \gtrless (1-\alpha)x_2 p_2(1-p_2)$，有

$$Y^P \gtrless Y^H \tag{1}$$

条件（1）有一个自然的解释。假设我们同时给项目做两项试验；两项试验结果为正的项目显然会被接受，那些两项试验结果为负的应被拒绝（然而，不存在做试验的理由）。这就产生我们在分离决策中应该做什么的问题。对好项目分离决策的概率是 $2p_1(1-p_1)$，以及对坏项目分离决策的概率是 $2p_2(1-p_2)$。因此，分离决策下项目的期望值是 $2[\alpha x_1 p_1(1-p_1)-(1-\alpha)x_2 p_2(1-p_2)]$。在分离决策下，如果项目的期望值为正，我们接受，如果为负，我们拒绝。但这正好是条件（1）。如果一个项目在分离决策下被接受，不存在第二轮评估的必要；多极组织优于等级组织。另一方面，如果一个项目在分离决策下被拒绝，那么有必要进行第二轮评估：等级组织优于多极组织。当坏项目对好项目的比例较大，以及当坏项目的损失相对于好项目的收益大时，第二轮评估有重要的价值。

我们的结论有另外一种解释，当 $\alpha x_1=(1-\alpha)x_2$ 时变得异常清晰。这时条件（1）与条件 $p_2 \leqslant 1-p_1$ 等价，这里 p_2 是接受坏项目的概率（第二类错误），$1-p_1$ 是拒绝好项目的概率（第一类错误）。在这个核心的例子中，多头组织是否优于等级组织，依赖于决策人犯两类错误的相对可能性（对这个问题更加一般的说明在我们 1984a 的文章中有论述）。

委员会：组织架构的另一种类型是委员会。委员会是由个体组成的集合并具备明确定义的规则以便进行决策（采纳一个项目），例如，多数原则或者全体一致。有 n 名成员的委员会，要求至少有 k 名成员同意采纳，接受一个项目的概率，个体对赞成评估的概率是 p，是

$$f^c(n,k,p) = \sum_{j=k}^{n} \binom{n}{j} p^j (1-p)^{n-j}$$

多头组织与只需委员会中的一个人批准项目即可执行的情况在形式上是相似的，而等级组织与需要委员会中的所有人批准项目才可执行的情况在形式上也是相似的。在我们 1984b 的文章中，我们探讨了委员会的最优规模和决策规则，我们还证明，例如，在 n 给定的情况下，产出是 k 的单峰函数；同时，对于对称情况（$\alpha x_1=(1-\alpha)x_2$），是否多或少于大多数意见仅仅依赖于是否 $p_2 \gtrless 1-p_1$。

二、复合组织

很显然的是，大多数的组织不是上面讨论过的一种单一形式（等级组

织、多极组织或者委员会），而是各种组织形式的复合体。我们可以以上述基本组织形式为基础，建立更加复杂的组织。例如，在一个等级体系中的多极组织，项目被批准的条件是至少获得其中一个决策单位的批准，但是每一个决策单位本身是一个等级组织。这有点像市场经济。如果每一个单位有 n 个等级，以及在一个多极体系中有 m 个类似单位，第 i 类项目被接受的概率是 $f(n,m,p_i) = 1 - (1 - p_i^n)^m$ 。通过适当地选择 n 和 m，我们可以确保 $f(n,m,p_1) > p_1$ 和 $f(n,m,p_2) < p_2$ ，也就是（n，m）多极—等级结构的组织与单一决策者相比，会接受更多好的项目和拒绝更多坏的项目。而且，它比简单多极组织（$n=1$）拒绝更多坏的项目，以及比简单等级组织（$m=1$）接受更多好的项目。我们可以建立更加复杂的组织，其中，上面所讨论的等级中的多极作为一个单位，以及第二层的等级中的多极用这样的单位来创建。只要具备足够多这样层次复杂的组织，我们就会实现完美甄别（perfect screening）。这是防止完美决策的决策成本。随着层次的增加，多层次组织的成本迅速增加，而且，增加的期望产出的收益却降低。两者之间的平衡实现组织的最优结构。

为什么具备足够的层次很容易实现完美甄别。等级可以成功地增加好项目对坏项目的比例，但是仅仅在撇掉许多好项目的成本之下。另一方面，在多极中，可考察的项目存货是“丰富的”。事实上，当多极中的单位数目增加时，任意一种项目被接受的概率趋近于1。我们所得结论的道理是，如果我们首先通过等级组织过滤项目集，然后通过多极组织，接受的项目集比通过单一过滤得到的更加“精练”，重复这样的过程会得到完美选择。

我们可以运用先前部分有关复合组织的分析来证明其他有趣的结论：例如，如果组织成本仅仅依赖于管理者的数目，那么，将一个特别长的等级组织重新安排进两个（或者多个）多极会更好；以及在一个等级组织中，将一个特别大的多极重新安排成两个（或者多个）多极子单位（参见我们1984b的文章）。

三、选择管理者：趋向组织动态

组织中最重要的决策是谁适合做什么工作。人们对这一问题的关注意味着情况会有所不同：个人在获得、交流和处理信息的能力上有差异。每一个组织结构都有对具有不同特征个体的最优安排。更为重要的是，一些组织形式的绩效与其他组织绩效相比，对这些安排以及组织架构自身功能安排中的误差，更为敏感。

组织选择继任者的“规则”的随机过程，描述了在组织中不同能力的个体被安排在不同的职位的过程。我们可以分析这些不同组织架构随机过程的稳态，以及比较它们的相对表现。我们将阐释这是如何实现的。

假设存在两类个体，有能力的（C）和无能力的（I）。假设在多极组织中的每一个人选择他自己的继任者，然而较高级别的等级的人选择他自己的继任者和下属。显而易见，系统中存在四个状态：（C，C）、（C，I）、（I，C）、（I，I）。我们可以说明 Pr（C，I）和 Pr（I，C）在多极组织中较大，然而 Pr（C，C）和 Pr（I，I）在等级组织中较大。如果两个体系的均值产出相同，任何风险厌恶社会将更喜欢多极组织（关于详细讨论，参见我们1984c 的文章）。

替代经济组织在选择纠错上的能力也有差异。因此，我们怀疑，一旦引入（展开的）消除有缺陷组织的机制（例如，市场体系中的破产），就会证明多极组织更好。

四、扩展

上面的分析忽略了与组织设计有关的成本决定的三个方面：（a）时间（组织层级越多，决策所需时间越多）；（b）交流成本（不仅有直接交流的成本，而且有在交流中不可避免出现的误差）；（c）决策的顺序（例如，在委员会中假设每一个人都审阅所有的项目，然而在等级中较高级别的人只审阅那些由较低级别送给他们的项目）。既然审阅有成本，审阅过程的顺序对评估所用资源产生巨大影响。

在其他论文中，我们已经研究过这些以及基本分析的扩展，包括（a）对贝叶斯决策规则运用结果的更加广泛的处理（即使在我们的基本模型中信息技术是给定的，我们的分析仍然是符合贝叶斯定理的）；（b）对信息获取的支出水平内生决定的分析；（c）替代组织形式对已有项目集所产生后果的分析（一个组织可用项目集是内生变量，由对开发项目者的激励决定，这反过来还部分地依赖于所采纳的不同类型项目的可能性，这在组织结构间有显著的不同）；（d）对面临不同问题的替代组织形式的评估，例如，选择项目最优集，而不是最大化期望收益；（e）外部性。例如，一个组织的决策，影响其他组织执行的项目的生产率。这种相互影响在特定情况下很重要，而且它们是可以强化等级的例证。

五、结论

不同组织结构产生误差的性质不同，这一点尽管重要，并不是组织形式中仅有的差异。我们没有考察经济结构与政治结构关系的普遍认识。例如，所谓的等级制度与独裁主义的关系。我们忽略了一些组织方面的比较，它们在文献中已经被广泛讨论。例如，传统的模型强调分权的计算优势（事实上，在一些版本中，组织形式的差异，似乎仅是解一般均衡问题时，替代运算法则间的比较）。经济体在“真实时间”解决问题的事实，而模型在虚拟的时间解出均衡，意味着（至少像传统展示的）这个争论仅是有限的相关。

我们也忽略了激励问题，而这一问题近来处在组织设计讨论的中心位置。我们认为激励问题十分重要，并且组织结构会对可能的激励集的决定产生重要影响。例如，当有一个以上决策单位时，人们可以依赖相对表现来进行奖励；当单位数目增加时，在特定情况下，可以实现一阶最优。（Barry Nalebuff and Stiglitz，1983）。然而，我们认为，组织可能运行得很差，不仅是因为误导的动机（就像激励文献中所阐述的那样），而且还源于人类易错性（fallibility）。我们甚至已经考虑说明，在激励问题缺失下，个体错误是如何在不同组织形式下进行不同的加总，导致组织运行的系统差异。

驳斥我们分析的一个观点是，等级制度永远可以自身分权，但是反之并不亦然；因此，似乎重复地说明，等级比多极形式更好。在这个方面，我们提出的问题是，在何种情况下，一个等级机构应该按多极结构来组织自己。但是我们认为，这多半有些误导：如果在等级结构中有干预的权力，当适当的情况产生时，该组织就有义务去干预，并且它还有必要获取信息，实现干预。仅当等级组织强有力地以及迅速地承诺不干预时，等级组织等价于多极组织。然而，这些问题的分析必须等待另外的契机。

本文中，我们大多数的分析基于替代经济系统的比较，但是我们的结论可以应用在经济分析的一些不同层面上（公司层面、产业层面或者整个经济体）。而且，我们的分析在政治决策背景中有直接和明显的含义，可以是组织的微观决策（委员会运行的规则，或者公共项目选择所执行的管理过程），也可以是国家组织。事实上，我们希望我们对长期存在组织的分析，对组织选择决策者所面临问题的分析，以及对组织绩效对选择过程性质的相对敏感性的分析，将有助于人们看清政治理论中关于不同政治结构优劣点的某些长期以来的错误。有关国家体系设计的传统讨论（Karl Popper，1950），基本上忽略了决策中产生的人类易错问题。例如，当柏拉图探讨贵族统治至

高无上时，他从来没有考虑过一段时间内选择贵族成员所产生的问题，或者选择失败的后果（今天，这些问题大家都很熟悉）。

在本文，我们的目的不是对经济体系的比较提出确定性的结论，而是鼓励大家换一个角度对经济学中最根本问题之一给予关注，以及阐述如果建立简单模型，为一些长期存在的有关不同经济体系相对价值的争论，提供一定的启迪。

参考文献

Lange, *Oskar and Taylor*, *Fred M.*, On the Economic Theory of Socialism, *New York*: *McGraw-Hill*, 1964.

Nalebuff, *Barry J*, *and Stiglitz*, *Joseph E.*, "*Information*, *Competition and Markets*," American Economic Review Proceedings, *May* 1983, 73, 278 - 83.

Plato, The Republic, *Oxford*: *Clarendon Press*, 1968.

Popper, *Karl*, The Open Society and Its Enemies, Part I—The Spell of Plato; Part II—The High Tide of Prophesy: Hegel, Marx and the Aftermath, *Princeton*: *Princeton University Press*, 1950.

Sah, *Raaj Kumar and Stiglitz*, *Joseph E.*, (1984*a*) "*The Architecture of Economic Systems*: *Hierarchies and Polyarchies*," *Working Paper No.* 1334, *National Bureau of Economic Research*, 1984.

______ *and* ______, (1984*b*) "*Economics of Committees*," *mimeo.*, *Princeton University*, 1984.

______ *and* ______, (1984*c*) "*Perpetuation and Self-Reproduction of Economic Systems*: *The Selection and Performance of Managers*," *Mimeo.*, *Princeton University*, 1984.

______ *and* ______, (1984*d*) "*Human Fallibility and Economic Organizations*: *The Architecture of Economic Systems*," *mimeo.*, *Yale University*, 1984.

信息与资本市场

信息与资本市场*

一、导言

股票市场是否就是个赌场？一心要在市场中取胜的人们，会不会花费了太多的资源去获取信息？真的有可能在市场中取胜吗？这几个问题的答案对于如何评价经济中股票市场的作用是非常重要的。按照新古典经济理论，股票市场是衔接消费者和厂商的一条重要渠道。原则上，股票市场应当为企业做出正确的投资决策提供信号。然而，倘若在极端的情况下，投资个体对于投资项目的价值以及可能发生的各种事件的相对概率一无所知，那么股票的价格就只会服从随机分布。从而股票市场就不能发挥资源配置的作用了。

关于股票市场中信息的作用，有两类截然相反的观点。第一种观点认为，要是股票市场能够良好运行，那么股票价格就能够反映真正价值。因而人们就不存在获取信息的激励，于是就不存在包含信息披露的市场均衡。就算股票价格不能够反映其真实价值，内部人的收益也是以牺牲外部人的利益为代价的，因此，理性的外部人就会拒绝再"玩"；如果真是如此，那么内部人也就失去了赚钱的工具了。

相反，赫舒拉发（Hirschleifer）提出，在纯粹的交换经济里，人们有着

---

* Joseph Stiglitz，"Information and Capital Markets"。本文是1971年12月在新奥尔良计量经济大会上的讲演文章的第二部分。当时关于此题目的讨论已经变得非常广泛。因此我无法一一致谢。我要感谢 M. Rothschild，C. von Weizsacker，R. Lindsey，V. Krishna，S. Salop，A. Dixit 以及 A. Weiss，同他们的许多讨论令我受益匪浅，我对于这个题目的兴趣则完全是受到 Paul Cootner 的启发，灵感是来自他本人和他的文章。他对于我的思考方式的影响以及所受他的帮助是显而易见的。

强烈的获取信息的动机；任何人只要拥有关于未来自然状态的信息（比如哪种证券会涨价、哪种会跌价），就能够通过扩大资本而盈利。依照这个思路，赫舒拉发给出了一个简单而又极富启发意义的例子。假设有两家厂商。倘若状态1发生，厂商1的产量为q，厂商2的产量为$q/2$；倘若状态2发生，厂商2的产量为q，厂商1的产量为$q/2$，首先假设两种状态的主观概率相等，如果人们选择某种资产组合是为了追求预期效用最大化，那么这两个厂商的相对价格就一定为1。但是，即使有个人知道状态1将要发生，那么他会大量购入厂商1的股票而卖空厂商2的股票。于是厂商1的股价就会飙升，而厂商2的股价就会下跌。知情的投机商就能够大赚一笔。

实际上，赫舒拉发认为，由于从获取信息得到的收益属于完全的利益再分配行为，因此任何发掘信息的行为都是没有社会价值的：充其量不过是损人而利己罢了。

这两种观点似乎是完全矛盾的：一种认为就算在获取信息方面的投资是有社会价值的，也可能出现信息的投资为零的状况；另一种认为就算在获取信息方面的投资是没有社会价值的，仍然会出现投资于信息的状况。我认为，这两种立场都是错误的——后者的错误在于隐含地假设了个体的行为存在非理性，或者非竞争性，或两者并存。前者的错误在于未能正确地认识到证券商在提供信息方面所扮演的角色，没有意识到市场扰动的存在会使得价格无法完美地将信息从知情的一方传达到不知情的一方。[1] 此外，前者还没有考虑到即使信息市场是自由进入的，均衡仍然可能是存在部分垄断特征的，个别人可能对于个别信息拥有着垄断力量。[2]

这两种观点都没有能够充分地区分出信息的私人价值和社会价值。很明显，理解资本市场是如何运行的核心在于分析：1）在市场经济中，个体获取信息的激励；2）价格能够在多大程度上反映个体拥有的信息；3）在企业经理的决策行为中，价格究竟扮演何种角色。

我认为，在一个完全竞争的市场经济里，获取信息的各种激励很少是具有社会价值的。实际上，在纯粹交换经济这样一个核心而又特殊的情况下，唯一的均衡里根本不存在花费在信息上的成本（也没有贸易存在）。在更加一般的情况下，当存在个体差异的时候，的确存在着获取信息的各种激励。

1 在 Grossman and Stiglitz（1976，1980a）以及 Grossman（1977）的文章里对此有着更加深入的讨论。

2 在 Stiglitz（1975a）的文章里对此有简要的论述。

但是如果信息确实是有成本的，那么市场价格就不能完美地反映个人所持有的信息。企业的市场价格不能完美地反映其真实价值有两个重要的后果。

首先，企业主兼经理人虽然知道企业的市值被低估，却无法很好地将其投资风险分散。他们控制下的企业更显示出风险规避的倾向，尽管产量（利润）与证券市场是零相关的，并且市场在为企业定价时也认为它们是风险中性的。我认为，企业并不是以追求市场价值最大为目标的，同时，也不受被兼并的威胁。其次，高质量的企业会试图向市场传递有关质量的信号。由于直接声明的置信力十分有限（比如宣布盈利的前景非常看好），市场往往采用间接的甄别手段。因为高质量企业的企业主兼经理人所持有的本企业股票的比例一定高于低质量的企业，于是内部持股比例可以作为一有效的甄别手段。[3] 然而，要想实现这一点，好企业的企业主兼经理人就必须持有高于在其他情况下所持有的内部股票份额。所以，就算是市场价格能够准确地反映出企业的利润，企业仍然表现出风险规避的倾向。但是还可能存在其他的情况，比如唯一的均衡中市场未能将好企业与坏企业甄别出来；此时为了不让企业主兼经理人的内部持股比例传递有关企业质量的信息，劣质企业就必须能够模仿优质企业，也就是说，它们也有相当大的内部持股比例。

其他的指标，比如资产负债表以及分红政策也都可以作为甄别手段。但在这里，我并不想分析这些其他的重要金融指标的含义。[4]

二、信息公开的证券市场存在吗

在本部分，我们更加详细地分析导言中所讨论的悖论。这里一共介绍了好几种关于这个悖论的解决方案。在第三部分，我们将只讨论其中的一种。

我们只限于考察纯粹的交换经济。假设每个人所拥有的对不同企业的产权就是个人禀赋，每个企业的产出在期末依照各人当期的产权比例进行分配。假设在期初与期末之间共有 T 个交易期，其间个人可以互相交易持有的股权。假设只有两个交易期，这一方面可以大大地简化分析，另一方面也不会失去一般性。在某一期的开始直到可以交易的那一时刻，个人可以花费其他资源来获取信息，比如各个企业的产量或者不同自然状态的客观概率。

假设每个人的偏好都相同，并且拥有相同的初始禀赋（他们对各个企

3　近来不完美信息的文献中有时称其为自选择机制的一个例子。而拥有的股份比例据说可以作为一个信号，参见 Stiglitz（1982），Spence（1973），以及 Salop and Salop（1976）的文章。

4　参见 Leland and Pyle（1977），Ross（1977），Stiglitz（1982）以及 Stiglitz and Weiss（1980，1981）的文章。

业拥有相等的股份）。假设各个企业在各种自然状态下的产量都是已知的，[5]但是各个自然状态的相对概率是未知的。同样，为了简化分析，假设只有两种自然状态。开始时所有的人都认为这两种状态是同等可能的。考虑下面的两种可能情况：

1）没有任何投资者研究了市场。那么在第一个交易日，市场上各个企业的相对价格就一定会反映出每个人先验的信息，以使得两种自然状态的概率相等。由于所有人都拥有相同的信息、相同的偏好、相同的禀赋，因此，实际中不会有任何交易发生。在两个交易日里的市场价格完全相等。

2）有一个投资者研究了市场。上面所描述的情况可能是个均衡吗？赫舒拉发就认为不是。因为可能会有一个人愿意用一部分资源来获取下一期真实状态的信息。假设在下一个交易日之前，他得到了这个信息。这样他就知道市场上哪些企业的价值被高估了，哪些企业的价值被低估了。于是他就会卖空那些价值被高估了的企业股票，用卖股票的收入来购买价值被低估了的企业股票。在两个交易日之间的间隔时间里，他就向公众披露了他的信息，这就抬高了被低估的企业的价格，打压了被高估的企业的价格，使得信息投资人的两部分交易都得到了资本回报，从而能够承担获取信息的成本。因此，对信息进行投资就是有利可图的。但是，由于在信息上的耗费并没有改变可以消费的物品的总量，因此，赫舒拉发把在信息上的耗费称为是社会浪费。一个人的得利是完全建立在其他人的损失之上的。并且由于获取信息是要耗费社会资源的，所以前者的得利要小于后者的损失。

我们至少可以从三个方面来反驳这个论断：

1）在这个论断中，没有信息的个体也太过天真无邪了。他们至少应该发现：如果他们能够维持原来的资产组合的话，会比介入到投机交易中更有益。换句话说，他们应该采取我所称为的非投机策略。这种策略所带来的期望收益，会高于赫舒拉发假设的那种天真无邪的策略——而且同样简单。

2）另一方面，可以反驳说，没有信息的个体比前面所设想的要复杂得多。他们观察到了市场的价格所传递的信息——当关于某个自然状态的证券价格被不断抬高的时候，那一定是由于有内部信息的人知道那个自然状态即将发生。这个特例同样意味着没有交易发生。前一个例子是因为人们拒绝交易，而在这里他们确实有交易证券的需求，但是市场均衡价格使得交易量为零。

5 实际上通常我们会把自然状态定义为对不同企业所有结果的一个完全描述。

3）最后，没有理由认为只有一个投资者会研究市场。考虑一下只有两个投资者会研究市场的特例。在第一个交易日之前，他们都知道了哪一个自然状态会发生。假设他们之间是竞争性的关系（不是合谋关系），他们就会互相竞买竞卖，抬高本来应该被低估的证券价格，打压本来应该被高估的证券价格。如此一来，等到下一个交易日的到来时，市场价格会和第一个交易日的价格完全相同。证券市场上这两个人既没有盈利也没有损失。相比没有研究市场的人，两个人都有损失，因为他们都进行了信息投资。

综上所述，赫舒拉发关于信息的过度投资的论点隐含了要求没有信息的人是非理性的，并且拥有信息的人是具有垄断力量的。这绝不是一个理性的完全竞争的证券市场。[6]

而悬而未决的问题是：是否存在一个有信息公开的理性的完全竞争的证券市场呢？我的观点是：1）如果每个人都拥有相同的禀赋、偏好和信息的话，那么在一个理性的纯粹交换经济里是不可能产生交易的；2）即使信息量的大小是内生的，一个理性的有信息披露的证券市场也无法存在。

三、纯粹的赌博（证券）市场

这里讲到的纯粹的赌博市场有两个主要性质：1）某个人的收益完全是以其他人的损失为代价的（这是个纯粹交换经济）；2）赌局的结果（所投注的事件）与参赌人的其他收入来源是完全不相关的。后一个条件将赌博市场与保险市场区别开来。之所以有交易（赌博）发生，是由于不同的人对于事件发生的概率分布拥有不同的看法。赌马可以视为纯粹的赌博市场的一个原型。至于是否应当把股票市场视为纯粹的赌博市场，这个问题将会在后面的部分加以讨论。

一个广为流传的看法是，一个风险厌恶的有理性的人是不应该（也不会）参加赌博的。大家都确信，即便某个人真的拥有某事件发生的信息也不应该去赌。因此，就算得到了关于各匹赛马的相对优劣状况的信息，这一

6　认为两个人不可能在同一时间获得同一信息的论点是无法解释这个问题的。由此倘若有人先得到了此信息，他就对此信息有垄断地位，那么市场就不再是完全竞争的了；如果有很多人都先后得到了此信息，那么每个人都是一个临时的垄断厂商——市场仍然不是传统意义上的完全竞争的，在两种情况下，如果赫舒拉发是正确的话，就要求没有信息的人在下面所定义的意义上是非理性的。

事实并不就会使赌马变得更加理性。由风险厌恶的理性人所构成的纯粹的赌博市场是不会有交易（下注赌博）发生的，也没有信息存在。这些论点如果是正确的话，对于我们理解股票市场的本质有着深远的意义。当然，前提是股票市场可以视为一个纯粹的赌博市场。

在本部分，我们将检验由风险厌恶的理性人所构成的纯粹的赌博市场的均衡的性质。这个分析的一个关键部分在于如何理解“理性”的概念。其中心问题就是怎样合理地假设一个人知道什么，他观察到了什么信息，在观察到的信息的基础上究竟是如何做推断的？

我们在这里所关注的是多期重复的市场，这个市场在很多方面都很像赛马。之所以人们普遍地认为赌马是非理性的，主要是来自于他们从赌马中得到的经验：“理性的人应当认识到他们永远没有赢的机会！”但是这种说法并不具有说服力，因为的确有人赢了，正是因为看到他们赢了，其他人才会留在这个市场里。然而，我却以为，那种先验地认为赌博是非理性的观点是正确的。我们的目标是阐明这个观点的确切含义。在以下的讨论中我们将考察理性的一些替代定义。开始，我们先分析被称为**弱理性**（weak rationality）的概念。如果行为人对盈利的预期值等于盈利的平均值，那么他就是弱理性的。

注意到这里的理性的含义要比通常在醒目的“理性预期”标示下的理性弱得多，那里的理性指的是建立在对盈利的先验分布基础上的预期与后验的盈利值相一致。

命题1：如果所有人都持有弱理性预期，并且所有人都是厌恶风险的，那么在纯粹赌博市场上不存在任何交易（下注）。

证明：为了简单起见，假设个人在赌博以外的收入是给定的。赌局的定义如下：如果某个特定事件发生，个人的收入为1；如果没有发生，他将要支付$\frac{p}{1-p}$，他选择下注为B^j为了

$$\max U(Y^j+B^j)\prod^j+U\left(Y^j-\frac{B^j p}{1-p}\right)(1-\prod^j) \tag{1}$$

$\prod^j$是第j个人关于事件的预期概率；于是有

$$\frac{U'(Y^j+B^j)\prod^j}{U'\left(Y^j-\frac{pB^j}{1-p}\right)(1-\prod^j)}=\frac{p}{1-p} \tag{2}$$

显然

$$B^j \gtreqless 0 \text{ 时有} \frac{\prod^j}{1-\prod^j} \gtreqless \frac{p}{1-p} \tag{3}$$

预期收益为

$$B^j\left[\prod^j - \frac{p}{1-p}(1-\prod^j)\right] \tag{4}$$

因此如果下注（$|B^j|>0$）一定存在着某些人的预期收益没有能够实现。此题得证。

另一方面，可以很容易地建立关于命题1 的逆命题：

命题2：如果市场上有些人表现出非理性，有些人表现出理性，那么在纯赌博模型里的均衡中必然有交易发生了（下注）。

证明：我们已经得到了

$$B^j \gtreqless 0 \text{ 时有} \frac{\prod^j}{1-\prod^j} \gtreqless \frac{p}{1-p} \tag{5}$$

令$\prod^*$为事件的真实概率。由假设可知，对于某些人，有

$$\prod^j = \prod^* \tag{6}$$

对其余人，有

$$\prod^k \neq \prod^* \tag{7}$$

因此，至少对两个人，有

$$\prod^j \neq \prod^k \tag{8}$$

即使不存在交易，那么对于所有人，有

$$\frac{\prod^j}{1-\prod^j} = \frac{p}{1-p} \tag{9}$$

这就与式（8）矛盾。

截至目前的讨论中，我们并未涉及对信息差异进行区分的后果，也没有涉及获取信息的动机。在那些最有趣的情形下（包括股票市场），个人都是可以拥有（或者能够获得）信息的，并且不同的人所拥有的信息是有差异的。看起来似乎正是这些差异使得市场得以运行。我们将会证明，即便个人拥有的信息是有差异的，只要他在下述意义上是理性的，那么在纯的赌博市场上就不会有交易发生。

然而，首先我们要检验一个有些不同理性的概念，在此种理性条件下，纯粹的赌博市场上会有交易产生。如果个人关于各个自然状态的平均预期值等于真实发生的概率，那么就称之为市场显示出**集体理性**（group rationality）。例如，假设罐子里装有红色和黑色的球，每个人都下注赌红球的比例。

每个人都可以从罐子中抽出 100 个球。那么这个样本中红球的比例就是罐子里的红球比例的一个无偏估计量。因此，个人对于红球比例的平均预期值等于其真实值。如果每个人都以各自的预期来下注，并且如果预期的分布是对称的，我们就可以证明如下的命题。

命题 3：对于发生概率接近 1/2 或者服从集中分布的事件，其市场赔率是对事件真实发生概率 1/2 的有偏估计。更一般地说，即使在大样本中，市场赔率也不会等于其真实概率。

证明：令 Π^* 表示事件发生的真实概率。考察预期概率与之相差正负 x 的人的投注的供求情况。按照对称分布的假设，持如此两种预期的人数是相等的。假设 $p=\Pi^*$。考虑第一种情况 $\Pi^*=\frac{1}{2}$。显然有 $B(-x)=-B(x)$，其中 $B(x)$ 是指预期概率为 $\frac{1}{2}+x$ 的人想下的赌注。如果

$$\frac{U'(Y+B)\left(\frac{1}{2}+x\right)}{U'(Y-B)\left(\frac{1}{2}-x\right)}=1=\frac{p}{1-p'} \tag{10a}$$

则有

$$\frac{U'(Y-B)\left(\frac{1}{2}-x\right)}{U'(Y+B)\left(\frac{1}{2}+x\right)}=1=\frac{p}{1-p} \tag{10b}$$

我们现在证明如果 $\Pi^*>\frac{1}{2}, B(x)<-B(-x), x>0$，有

$$dln\left\{\frac{U'(Y+B)(p+x)}{U'\left(Y-\frac{pB}{1-p}\right)(1-p-x)}\Bigg/\frac{U'(Y-B)(p-x)}{U'\left(Y+\frac{pB}{1-p}\right)(1-p+x)}\right\}\Bigg/Y-\frac{pB}{1-p}\Bigg/dp$$
$$=\frac{2\times(2p-1)x}{(p^2-x^2)((1-p)^2-x^2)}+\frac{B}{(1-p)^2}\left[\frac{U''\left(Y-\frac{Bp}{1-p}\right)}{U'\left(Y-\frac{Bp}{1-p}\right)}+\frac{U''\left(Y+\frac{Bp}{1-p}\right)}{U'\left(Y+\frac{Bp}{1-p}\right)}\right] \tag{11}$$

在 $\Pi^*=\frac{1}{2}$ 这一点，对于所有 $B>0$ 它都是负的。因此对于 Π^* 稍稍大于 1/2 的情况，当 $p=\Pi^*$ 有

$$\frac{U'(Y-B)(p-x)}{U'\left(Y+\frac{pB}{1-p}\right)(1-p+x)}=\frac{p}{1-p'}$$
$$\frac{U'(Y+B)(p+x)}{U'\left(Y-\frac{Bp}{1-p}\right)(1-p-x)}<\frac{p}{1-p} \quad (12)$$

于是有 $B(x) < -B(-x)$。因此，对于$\int B(x)dx=0$，$p<\prod^*$。

同样地，如果 $x=0$，$B=0$，那么式（11）的右边部分就等于零。对 x 求偏导数，得到当 $x=0$ 时，（因为$\frac{dB^j}{dx}=\frac{dB^j}{d\prod^j}=1/2\prod^j(1-\prod^j)A(Y)$这里 $A=-U''/U'$）有

$$\frac{2(2p-1)}{p^2(1-p)^2}-\frac{1}{p(1-p)^3}<0 \quad (13)$$

对于$\frac{1}{2}<p\leqslant 1$，同样地有 $B(x) < -B(-x)$。

当$\left|\prod^*-\frac{1}{2}\right|$很大,并且 x 也很大的时候,市场上的赔率既可以是偏高的,也可以是偏低的。[7]

当然,在任何市场上,只要事件的市场赔率不等于其真实概率,就会刺激人们去获取真实概率的准确信息,当然前提是信息不能太昂贵。就算是在刚才所形容的市场表现出集体理性的条件下,难道某个人居然会傻到想不到其他人是依照样本的信息来下注的？人们的下注行为理应传递信息,而且理性的个体形成预期时理应将它们纳入其中。

如果在重复赌局中,有很多次都出现个人所能够观察到的所有信息的预期收益等于平均实际收益的情况,我们就称之为个人是理性的。这个概念同先前所述的弱理性概念是不同的,前者要求个人预期要建立在所能够得到的所有信息之上。但是同弱理性的概念一样,相对于传统的理性预期而言,这个概念也是一个条件相当弱的概念。因为它只是要求预期分布的均值等于真实分布的均值而已。由于每个人所观察到的信号都不同,他们拥有的信息也会

7　令 $U'(Y)=e^{-Y}$，令$\prod^*=\frac{3}{4}$,$x=\pm\frac{1}{8}$,那么对于$x=\frac{1}{8}$，如果$p/(1-p)=3$，由式（2）可知$\left(\exp\frac{-B}{1-p}\right)\cdot 7=3$或者$B=\frac{1}{4}\ln\frac{7}{3}$同时对于$x=-\frac{1}{8}$，有$\exp\frac{|B|}{1-p}\frac{5}{3}=3$或者$|B|=\frac{1}{4}\ln\frac{9}{5}<\frac{1}{4}\ln\frac{7}{3}$。于是有$p>\frac{3}{4}$。

不同，因而，会形成不同的信念。人们也许会猜想在这个条件下，赌博市场上可能会存在交易。我们现在证明不是如此。

命题4：如果所有人都持理性预期，在纯粹的赌博市场上就不会有交易发生。

证明：为了简单起见，不妨假设有两群人。第一群人只能观察到市场的赔率。第二群人能够观察到与下注的事件相关的信息。前者我们称之为没有信息的人，后者我们称之为拥有信息的人。由于没有信息的人只能观察到市场的赔率，如果他们参与交易，利用式(2)，一定会有

$$E\{\prod^{u}|p\}\neq p \tag{14}$$

这里上标 u 表示这是没有信息的人的预期值。拥有信息的人将其预期建立在某些信号 Ω 上，也可能还有 p。显然，如果他们忽略掉信号 Ω，那么他们的预期效用必然小于或等于不忽略此信息的情况的预期效用，

$$\max_{(B)}\mathrm{E}\{U^{j}|\ \Omega,p\}\ >\ \max_{(B)}\mathrm{E}\{U^{j}|\ p\} \tag{15}$$

假设 B^{*}（Ω，p）是拥有信息的人的最优交易函数，且对于某些 Ω 值有 $|B^{*}$（Ω，p）$|>0$。

但是，比如说在　$E\{\prod^{u}|p\}>p$ 且 $B^{u}>0$　(16)

条件下，显然 $B^{j}<0$，对 p 的期望盈利就是负的了。另外，按照假设，信号以及事件的发生都与个人收入无关。这样赌博的平均盈利就随着收入分布的更加分散而降低，从而预期效用也降低了。

$$E\{E\{U^{j}|\Omega,p\}|p\}\ <E\{U^{j}|p\} \tag{17}$$

这就与式（15）矛盾。因此，必定有 B^{*}（Ω，p）$\equiv 0$。

这个结果可以很容易地扩展到多于两组人群的情况。其背后的经济学直觉是很简单的：赌博不能增加各组人群的平均收入，而且至少有一组人群的收入是降低了的。与此同时，赌博增加了收入分布的分散程度。因此，至少有一组人群的期望效用是一定会由于赌博而降低的。由此可见，理性的人群中是不会存在赌博市场的。

这里所做的论证又可以同样扩展到 γ 不为零以及个人禀赋不相同的情况。纯粹的赌博市场的一个关键性质就是不能够发挥任何的保险功能，即如果对发生自然状态 S 下注，对于所有 i，j，有：

$$\frac{EU'^{i}[Y^{i}|S]}{EU'^{i}[Y^{i}|-S]}=\frac{EU'^{j}[Y^{j}|S]}{EU'^{j}[Y^{j}|-S]'} \tag{18}$$

$-S$ 代表所有其他 S 不发生的自然状态。在信息完美的条件下，市场上不会

存在对 S 的下注（原先关于自然状态 S 的资源配置在事前已经达到了帕累托最优）。假设有些人拥有信息 Ω 满足

$$\frac{E\{U^{i\prime}[Y^i|\ \Omega,S]\}}{E\{U^{i\prime}[Y^i|\ \Omega,\ -S]\}}=\frac{E\{U^{j\prime}[Y^j|\ \Omega,S]\}}{E\{U^{j\prime}[Y^j|\ \Omega,\ -S]\}} \tag{19}$$

这就说明，假如每个人都已经拥有了这个信息，就不会有下一步交易发生；假设某一时刻只有一部分人拥有这个信息。关于“零赌博市场”的定理显示仍然不会有交易发生。由于 B（p）>0，对于没有信息的人，有

$$\frac{EU'^u[Y|p,S]}{EU'^u[Y|p,\sim S]}\frac{\prod(p,S)}{1-\prod(p,S)}>\frac{p}{1-p} \tag{20}$$

$\prod(p,S)$是条件于所观察到的市场赔率的 p 所发生的概率，而它只依赖于 Ω。

假设 p 并不依赖于 Ω。只有当

$$\frac{EU'^j[Y^j|S]\prod(p,S)}{EU'^j[Y^j|-S](1-\prod(p,S))}<\frac{p}{1-p'} \tag{21}$$

时，拥有信息的人才会“出售”赌局，然而，利用式（18）和式（20）这显然是不可能的。

再假设 p 确实依赖于 Ω。为了使得拥有信息的人愿意“出售”赌局，必须有

$$\frac{EU'^j[Y^j|S,p,\Omega]\prod(p,S,\Omega)}{EU'^j[Y^j|\sim S,p,\Omega](1-\prod(p,S,\Omega))}<\frac{p}{1-p'} \tag{22}$$

对于所有 Ω 以及给定的 p 都成立。由于 p 是 Ω 的函数

$$\frac{EU'^j[Y^j|S,\Omega]\prod(\Omega,S)}{EU'^j[Y^j|\sim S,\Omega](1-\prod(\Omega,S))}<\frac{p}{1-p} \tag{23}$$

对于所有满足

$$p(\Omega)=p \tag{24}$$

的 Ω，此式一定成立。但是对于每一个 Ω，只要式（24）成立，必然有式（19）成立。于是有

$$EU'^u[Y^u|S,\Omega]\prod(\Omega,S)<\frac{p}{1-p}EU'^u[Y^u|\sim S,\Omega](1-\prod(\Omega,S))。 \tag{25}$$

对于满足式（24）的所有 Ω 作预期函数，得到

$$EU'^u[Y^u|S,p]\prod(p,S)<\frac{p}{1-p}EU'^u[Y^u|\sim S,p][1-\prod(p,S)] \tag{26}$$

这与式（20）矛盾！

从直觉也可以这样来表达这个结果：任何没有信息并且也知道自己没有信息的人都不会下注，因为他知道对手只会在赔率对自己有利的情况下才会下注。假设他并不了解自己是否比对手知道得更多或者更少，但是却了解到对手很清楚是否比自己知道得更多或更少。那么对手就只会在比自己知道得更多并且未来对于对手自己有利的情况下才下注。由于每个人都知道这些，因此他们就不会愿意参赌。于是唯一可能存在赌博的情况就是双方都不甚了解自己是否知道得更多；这样他知道得多的概率就同知道得少的概率相等。但是如果这样的话，下注不过是将收入"期望保持扩展"了，如果双方都是风险规避的（无论他们的效用函数如何），他们都不会愿意参赌。[8]

四、在纯粹赌博市场里获取信息的激励

命题4有一个很重要的推论。因为在纯粹赌博市场里无论是否存在信息的差别，只要人们持有理性预期，就不会有交易存在，人们就永远不会有获取信息的激励。如果我们真的在市场上观察到了交易发生，那么我们就可以推断要么这个市场上的一些人不是理性的，要么这就不是一个纯粹赌博市场。如果个人在不同自然状态之间的边际替代率是不同的，交易就可能是合意的。显然，保险市场是可以实现这个功能的。尽管股票市场上的某些交易也可以实现这个功能，但是这是否能够成为交易产生的最初动机，确实值得怀疑。

即使是在存在交易的纯粹交易市场上，还是存在着同样的问题：是否有激励个人获取信息的因素存在。有一种看法认为，既然存在着交易，那么只要可以获得更多有关真实概率的信息，比方说这两种自然状态的概率，就能够提高期望效用（前提是获取信息的成本要足够小）。但是这是不正确的。因为如果有人拥有信息的话，那么市场价格就有可能传递全部的信息。然而如果市场价格已经传递了全部的信息，那么再去投资信息就没有意义了。

在价格完全传递信息的条件下，存在信息成本的一般均衡

一个简单的例子可以更清楚地说明这种情况。[9] 假设有针对下雨的保险市场。是否真的会下雨取决于某些可以观察到的因素，但是却要付钱才能观

8　最近 Milgrom and Stokey（1979）的著作利用 Aumann 的共同知识观念（1976）将结果一般化了，关于分布的期望保持展型，参见 Rothschild and Stiglitz（1976）。

察到。如果不付钱的话，只能观察到一部分因素，其余的观察不到，于是就无法准确地预测是否会下雨。令 η 代表信息，$\prod$ 代表在 η 信息下的预期下雨的概率：$\prod = f(\eta)$。简单起见，假设 $f' > 0$，且假设每个人事前所观察到的天气情况都一样。对“天气下注的净需求”将会是预期下雨的概率的函数（而且还是每注的价格，即赔率的函数）。令 p 代表这个价格。有，$B_i = B_i(p, \prod)$，B_i 代表拥有信息的人群对下注的人均需求。价格上升降低需求，而事件发生的预期概率会增加需求。因此有，

$$\frac{\partial B_i}{\partial p} < 0, \frac{\partial B_i}{\partial \prod} > 0 \tag{27}$$

而没有信息的人（他们看不见 η）只能依赖于先验信息（假设给定并且始终不变）以及市场价格（赔率），因为每个人都知道，如果存在拥有信息的人，那么市场赔率一定会将他们的信息反映出来。若我们以 $B_u = B_u(p)$ 来表示一个人的下注需求函数 $\partial B_u / \partial p < 0$，那么市场出清则要求有 $B_i(p, \prod) + B_u(p) = 0$。在我们的假设下，对于每一个 $\prod$ 值都会有唯一的 p 值与之对应：就是说，没有信息的人完全可以从市场赔率中推断出来拥有信息的人的全部信息（尽管市场赔率不必等于）。这样就不会有人愿意获取信息了。注意到如果没有人获取信息，那么市场赔率在各期就会保持不变，因而就无法反映任何信息了。此时，如果信息的成本不是太高的话，那么获取信息的行为就又会变得有利可图了。

因此，在信息成本存在而又不是太高的情况下，纯粹交易市场是不存在非随机的一般均衡的。如果下注的需求函数有赖于其他不可观测的随机变量（并且它们的值也不能直接地由市场中的其他变量推断出来），那么市场价格（赔率）也会依赖于这些变量的，这样没有信息的人就不能从市场赔率中准确地推断出信息了。因此，在非纯粹赌博市场里，可能存在着交易。当然，此时价格就不能完美地反映出人们所拥有的信息了，尽管在平均的意义上还是可以的。在这个体系中，没有信息的人们会被激励去研究拥有信息的人以及别的没有信息的人的需求函数，而不是仅仅从 η 直接地推断，因为一旦知道了需求函数和市场价格，他们就能够推断出来关于 η 的信息，这可能是没有信息的人的一种（实际上是唯一的）办法，得以廉价地获取那些拥有信息的人所能够获得的全部信息。

9　这样的例子在 Grossman（1977）以及 Grossman and Stiglitz（1976，1980a）的一系列文章中都得到充分的扩展。

五、作为赌场的股票市场

很显然，股票市场并不是纯粹的赌博市场。但是，在接下来的部分里，我们假设市场上的大部分证券交易都是建立在某种非理性的基础上的，并且只有很有限的激励因素，去促使人们获取有关不同证券的回报的信息。

为了简单起见，假设所有人的工资收入都是常数，并且初始时刻对每个企业的证券所拥有的股份份额都相等，而且是风险厌恶的。一开始我们假设所有人的风险偏好程度都一样，我们对非投机者（nonspeculator）的定义是：无论市场价格如何，他都要购买（保持）一个代表性的股份。而一个投机者（speculator）的定义是：他会先建立对各种证券回报的预期，在给定这些预期以及市场价格的条件下，优化其证券投资以寻求最大化其期望效用。个人所形成的预期回报可以包括所有证券，也可能只是包含一部分的证券。

第 j 个人对于第 i 个证券回报的预期若满足下式就称为是无偏的（其中~符号表示真实值，^符号表示估计值）

$$E\,\bar{\mu}_i^j = \hat{\mu}_i, E\tilde{\sigma}_{ik}^j = \hat{\sigma}_{ik}^j \tag{28}$$

$\hat{\mu}_i^j$ 代表第 j 个人对于第 i 个证券的预期平均回报。σ_{ik}^j 代表他对于此证券与第 k 个证券回报的协方差。

命题5：*如果让人们在付钱获得无偏的信息以及遵循非投机策略之间选择的话，所有人一定会选择及遵循非投机策略。*

假设只有一个人进行了投机。显然，假如其他人都遵循非投机策略，那么他持有各个企业的股份一定是那个代表性的股份额。因此他的收入就同别人一样，但是由于他在信息上耗费了资源，所以他的福利降低了。如果投机的人不止有一个，而且他们获取的信息都一样，那么同理可以证明结果是一样的；如果他们获取的信息不一样（但是总的说来是无偏的），那么他们一定可以得到市场平均回报，当然他们有时候得到的回报高于市场平均值，有时候要低于市场平均值。如果信息的质量同这个经济的产出不相关（于是在市场表现得差强人意之时，个人之所得不会系统地高于平均值，反之亦然），那么个人的收入的方差就会高于不投机的情况，并且与前面所述的情况一样，他的平均净收入会降低。投机同样地是无利可图的。

即便是在有人具有获取信息的比较优势的情况下，这里所得到的结果依然成立。例如，假如某个人可以没有成本地获取不同证券回报的确切信息，

那么其余的人就不会愿意和他交易。这样他就仅仅只是决定了各个证券的价格而已。假如有两个人都具有获取信息的比较优势——并且两个人得到的各个证券回报率的信息都一致 ——同样地，如果获取信息，他们就会得到负的净收入。实际上，这是缘于他们的最终资产组合会和与执行非投机策略的情况相同。

这个结果可以很容易地扩展到人们具有不同效用函数的情况下。假设人们对于各个证券的评价依赖于其所带来收入的期望与方差。如果他们必须选择：1）获取每种证券回报的无偏的信息；2）采取非投机策略；3）对于风险资产采取非投机策略，但是同时也获取每种证券回报的均值与方差的信息，但是相比获取每个证券回报的信息的成本要小。我的观点正是：他们会选择后两者之一。

在每一期都获得无偏而且可信的信息的人总会购买相同的证券资产组合——市场风险资产组合。而仅仅获取市场风险资产组合的期望与方差的信息的人将会获得关于“市场曲线”信息，[10]其准确度与拥有每种证券回报信息的人的一样。但是由于他所付的信息成本较小，因而他的福利更高。

因此，在均值—方差模型里，人们永远不会去获取每个证券回报的信息。[11]

注意，一般来说，对于股票真实价值的信息进行甄别（将于第七部分讨论）将不会从整体上改变个人对于市场的信念。大部分由股票交易人所提供的信息都是这个类型。

六、关于股票市场上交易与信息获取的另外两种解释

然而，股票市场上确实存在着交易，而为了得到关于某种证券的信息，人们并不吝啬。怎样把这些现象同先前的分析融合起来呢？关于交易的两种基本解释是：1）在我们所定义的意义上，有人不是理性的；2）人们的禀赋和偏好（包括对风险的态度）不同，因此就算信息是相同的，交易也是合意的（由此或许可以引申出：需要改变他们的资产禀赋，包括不可交易

10　在所谓的期望方差分析里，市场线由下式定义：$\sigma_y = b(\mu_y - r)$，而μ_y和σ_y是由市场上所有风险资产组成的资产组合收益的期望和标准差，r是无风险利率。

11　此结果可以视做命题4的直接推论。在期望—方差模型里，关于期望和方差的信息是不会改变预期的，而且也不会改变个人的资产组合配置。于是式（18）和式（19）都满足；更加详细的讨论见 Grossman（1977）。

的资产，直到确保股票市场上的交易是连续的)。我们将在下面两个部分探讨这另外两种基本解释。

1. 非理性

这一点要证明只要有连续形式的新的个体诞生，就可能存在投机行为。这个理论要求每时每刻都必须有傻瓜出生，但是要求我们每个人都必须相信他不是傻瓜。[12]

当某个人进入市场看见了投机的收入分布，假如他觉得自己是普通投资者，那么他是不会投机的。然而假如他认为之所以有人的收入高于普通人，是因为他们具有获取信息的比较优势（不是仅仅因为运气好——另一种假设，认为数据面前人人平等)，而且他觉得自己就是这样的人，那么他就会毫不犹豫地投机。如果存在很多这样的人，那么就会形成一个“竞争性”的证券市场。对于那些“获胜”的人而言，他们更加认为自己具有比较优势；而那些“失败”的人会逐渐觉得自己没有比较优势。当他们失败的次数足够多了，他们就会停止投机交易而采取非投机策略。

市场的“厚度”取决于新的市场进入者的流量、人们调整预期的速度以及证券回报的方差。假设他们都很笨，而且调整预期的速度都很慢（而且这些赌徒的好运就要到头了)，那么他们只有在一连串的失败过后才肯退出市场。

2. 禀赋和偏好的差异

股票市场上的一些交易贯穿着一生——比如，年轻人购买证券然后在年老时出售。但是如果前面的分析是正确的话，那么理性的人们就应该去购买共同基金，而不是交易各人的证券。

如果企业家很幸运，创建的企业的业绩斐然，那么本企业的财产占他们所有财富的绝大部分。假如他们是风险规避的人，他们就会希望出售部分企业股份，购买包含整个市场的共同基金。这样就会产生交易；然而，如果前面的分析是正确的话，那么交易是不会持续下去的。一旦个人的资产组合达到“平衡”，个人持有的各个股份比例刚好，交易就会停止了。

并不是所有的资产都是可以交易的，人力资本市场就是一个众人皆知的不完美市场。由于第一，不同人的人力资本收益之间是有相关性的；第二，市场上不同的证券对于不同的人力资本回报的相关性也是不同的，如

12 这一点对应于如下的现象：很多老师观察到超过一半的学生认为自己处于中等偏上的位置。

果人们在开始时的资产禀赋相同，那么他们就会有交易的动机。其目的为了使资产组合与人力资本相匹配。一般而言，人力资本的变化会使得最优资产组合发生变化。因此，矿工会卖空期资产组合里过多的矿产（假设作为矿工的技术是该产业特有的）。但是我怀疑这种证券交易的“保险”或者“匹配”功能是否真的重要。在任何情况下，假如这是原始动机的话，那么它对于信息获取的激励性质会产生非常重要而且有趣的影响，这一点将在下一节讨论。

关于人们只是由回报的均值和方差来评估资产的这一假设，对于只有与市场资产组合相关的信息才是有价值的结论是十分关键的（当人们对于风险的偏好不同时，关于市场资产组合回报的均值和方差就具有了社会价值，因为他们能够使得风险的承担变得更有效率，正式地讲，对于这样的信息，式（18）是不成立的）。当人们对于风险的偏好不能由资产组合回报的均值和方差来描述时——（或者由满足一般共同基金定理的另外一种效用函数形式描述（Cass and Stiglitz，1970）——此时，关于某个证券的信息就又是有价值的了（同样地，式（18）是不成立的）。

但是要注意如果所有证券的回报可以写为一组市场要素的线性函数，

$$Q_i = \sum M_j \mu_j + \epsilon_j \tag{29}$$

当证券的数量足够多，多到个人可以将通常某一证券所特有的个别风险 ϵ_j 完全分散掉的程度，此时，关于某个证券的信息就又变得没有价值了。个人的预期效用就是每个市场要素的影子价格（假设小企业里他不受信息的影响）以及这些市场要素的分布的函数。因此，事后得到的效用水平仅仅是向量 M 的值以及各个市场要素的影子价格的函数（当然，还包括这些市场要素的分布）。同样，如果式（18）是成立的，那么式（16）也是成立的；任何证券的信息都不会导致交易的发生，于是，这些信息是没有价值的。

七、信息的回报

这一部分，我们将要更加详细地检验在股票市场上获取信息激励的性质。我们将会证明信息的社会价值与私人价值之间是有很大的区别的。不过，首先我们必须分辨出几种不同的信息。

1. 信息与资本市场的分类

在我的另一篇关于信息和劳动市场的（Stiglitz，1975b）文章中，我区分了两种信息：一般信息与特别信息。前者包括一个人的个人特征，它可以

影响很多种职业的产出，后者是指与特定职业相关的信息。

根据信息与交易双方的关系程度，这里我们需要分类出四种信息。

1）一般——一般：影响所有的证券以及所有买家（比如，各种自然状态的相对概率本身所固有的信息，它对所有人都同样重要，例如衰退发生的概率）。

2）一般—特殊：只影响特定的人，但是会改变他们对一整类证券的态度（比如，采矿业发生萧条的概率对于从事采矿业的人而言是非常重要的，但是如果采矿业的衰退同整个经济周期不相关的话，那么它对其他产业的人就不那么重要了）。

3）特殊——一般：影响所有该种证券的持有人，但是只影响某一个企业（比如，经理的思想是否稳定）。

4）特殊—特殊：只是对个别人（或者一类人）以及个别企业有价值。倘若我是风险中性的人，那么我就只关心我的证券的平均收益情况。倘若我是风险厌恶的人，那么我也只关心我的证券收益风险的其他情况 。

大多数信息都可以在非常一般与非常特殊之间找到定位——一般地，信息不会影响到所有企业和个人，但是往往会对不止一个企业和个人有价值。然而，我们对信息所做的四种区分是极其重要的，因为获取信息的收益机制以及个人与社会收益的关系会随之不同。

我们还必须做一项重要的区分：有些信息对不同人都有价值，但是对于某些人可能会使得其证券升值，而对于另一些人，则可能导致其证券贬值。

如果某种意外事件发生会使得企业表现很好，而不发生则会使得企业表现很差。这种信息可能会让确信事件发生的人高估该种证券 ，而不相信事件发生的人则会低估该种证券 。[13]例如，一个矿工的工资总是同采矿业的产业特点相关的，因此他对那些盈利与采矿业整体表现负相关企业的评价会高于那些与之正相关的企业；而对于另一个产业的人来说，其所在产业与采矿业负相关对应着与他的工资正相关，因此，他也许会觉得这个证券缺乏吸引力。

我们略微考察一下那种可以影响两种事先无法区分证券的信息。如果信息使得所有人都最终同意证券 A 比证券 B 要值钱，那么就把此信息称为*分级的*。如果信息使得有些人认为证券 A 比证券 B 要值钱，而另一些人认为证券 B 比证券 A 要值钱，那么就称此信息为*匹配的*——它把特定的人同特

---

13 显然假如存在着全套的阿罗—德布鲁证券市场，那么各种事件发生的相对概率就可能会影响到各种的连续商品的相对价格。但是给定阿罗—德布鲁证券的价格，企业就无需考虑不同事件发生的相对概率了。见 Stiglitz（1970）以及 Grossman and Stiglitz（1977，1980b）。

定的证券匹配起来了。[14]

2. 交换经济中信息的社会价值

从前面的论证中可以发现交换经济中信息的社会价值的本质。[15]只要个人是异质的，那么一般而言，每个人都持有相同的资产组合就不是最优的。信息（无论是一般还是特殊的）使得证券在不同的人之间得到更好的匹配。

3. 交换经济中信息的私人价值

提供信息的私人回报有两种形式：1）直接的消费效应，尤其是“特殊—特殊”的情况，因为它使得个人可以通过获取信息来优化资产组合以满足需要；2）市场的价值效应；倘若公布信息会使得证券的市场价值发生改变的话，那么通过卖空高估了的证券，买空低估了的证券，个人就能够通过公布信息来赚上一笔。在这一讨论点，我们重点考察后一种效应，尤其关注人们获取并且披露信息的激励因素。[16]为了使得个人能够独占（appropriate）这一类信息的收益，要求：

1）在信息被披露之前，他们必须拥有或者能够获得资产；

2）一定存在有人披露信息的激励因素，使得市场价格可以通过调整来反映信息；

3）信息一旦披露就必须是可信的。

然而，这三步分析中每一步都会遇到问题。在前面的分析里（第二部分和第四部分）证明了纯粹赌博市场中是没有交易的。实际上，要是有人（假设他拥有信息）愿意购买我的一些股份，那么就传递了这样的信息：他认为我的股票被低估了。于是我不会将股份卖给他。一个很有趣的引申含义是：通常如果企业的市值被低估，那么它会被收购掉。但是如果被收购企业是像上面那样想的，收购就永远不会发生了（只要在平均意义上，被收购企业的市值确实是被低估了的）。[17]

信息披露会产生标准的公共经济学的问题（搭便车）。每个股东都希望别人来支付获取信息的成本，而所有持有股份的人都将从股票价格上涨中获

14 这个区别在我对于劳动市场的信息分析中占有重要地位。某个企业的薪酬信息是分等级的；关于非特有特征的信息则很可能是匹配的。

15 当然，引入生产后的收益更大了。

16 等级信息会影响所有投资者对于资产的评价；但是不会使得资产得到更好的配置。

17 当然，兼并的动机还有许多；例如有钱人会购买同他的其他资产收益负相关的股票。如前所述，交易可以使得资产在人群间更好地匹配，但是它很难成为兼并的动机。

得好处。[18]只有当个人占有所有股份或者他们合谋行动，才可能使得信息的提供方攫取所有的由于股价上升带来的收益。[19]

但是，企业主却可以很自然地采取一种合谋策略：拥有本企业的信息。假设企业主拥有所有关于本企业业绩的信息（这是一致的），那么显然只要信息传递成本不是太高，企业就会受到披露信息的激励。

业绩不佳的企业当然会隐瞒信息。但是如果只有一部分企业披露了信息，那么市场就会认为那些不披露信息的企业是业绩不佳的。由于没有关于它们的任何信息，所有这样的企业都会被一视同仁。[18]然而对于那些业绩好于这些差企业的平均值的企业，它们可以通过有意或无意地披露自身信息来增加企业的市值（当然这是以牺牲那些业绩不佳的企业的市值为代价的）。这个过程一直重复直到每个企业的信息都被披露为止。这个过程与我曾经使用过的教育系统中的甄别机制模型是紧密相关的。在那里，聪明人得到大家的承认是符合他自身的利益的，尽管获得承认的过程是没有社会价值的（如果这样不会导致产出的变化的话），因为他得到的回报是以牺牲了蠢人的利益为代价的。

但是，资本市场上的甄别机制要比教育市场上的情况复杂得多，主要是因为有几个重要的区别：最重要的区别来源于人力资本是无法拿来买空卖空的（至少目前是如此）。也就是说，即使我知道你的价值被高估了，我也没有办法和你签一个合同，将你的服务卖到下一期去，从而公布你的真实生产能力，而后等到下一期以低得多的价格再买回你的服务。我也不能只买你一半的人力资本，然后留着以后再卖。[20]因此，我们在讨论甄别模型的时候，是假设个人自己选择是否把自己甄别出来的。由于资本市场里每个人都可以是某个企业的老板，通过投机交易，至少存在着这样一种可能性，就是可能存在企业主以外的某个人，他同样希望公布该企业的特征信息。但我们在分

18 为了改善管理绩效而兼并同样会存在着搭便车的问题。平均而言，兼并的确可以改善管理绩效，会提升资产的收益，从而能够提高市值。但是小企业的股东不会受到出售股份的激励，他可以从改善管理绩效获得好处。只有当新的管理可以为自己带来资本收益时，兼并才有可能发生。Grossman and Hart（1980）对此有进一步论述。

19 注意到如果存在着很大的信息传递成本，而且该成本是信息传递人数的函数，那么唯一可能从传递信息中获利的人，就是其他潜在的大股东（有钱人）。因此，有钱的卖家会向有钱人提供证券的信息。

20 这种说法不太正确，企业的确希望能够找到“被低估”的人，将他们区分出来，然后以更高的价格出售他们的服务，这等价于出售一部分他们的人力资本。但是这只是例外而绝不是一般规则。

析时假设甄别的收益完全由证券的原始持有人获得——同教育市场一样——只有他们才有提供信息的激励。

在教育市场中，我们通常假设个人对自己的能力比对别人的能力更加了解（当然这不是必要的）。同样地，在资本市场条件下，一般建设证券的原始持有人拥有该证券的最多信息。[21]

4. 信息的私人价值与社会价值之间的关系

前面我们已经证明了资本市场上的某些问题可以由甄别的经济理论重新阐释。这样我们就可以借用甄别理论的一些一般结论：

1）提供信息的回报主要被证券的原始持有人获得，这种信息改变了企业的市值，因此只有他们才有提供信息的经济激励。

2）信息的私人价值与社会价值一般是不相等的，正如教育市场上甄别的例子那样。对于“分级信息”有可能存在着投资过度——即甄别出来的信息对于每个人的价值都一样——而对于“匹配信息”有可能存在着投资不足——即甄别出来的信息对于每个人的价值不一样。后者可以很容易地通过一个例子说明：假设披露信息不会改变资产的价值，那么某一类人对于证券的需求上升一定会被另一类人的需求下降所抵消。即使提供信息没有任何成本，并且就算提供信息具有很高的社会价值，企业也没有激励去提供信息。

下面的两个部分，我们将会给出一些例子来说明在分级信息和匹配信息的情况下，信息的私人价值与社会价值之间的关系。

上面所提的第三个问题是置信力问题：尽管企业受到提供信息的激励，但是他们同样受到提供虚假的信息的激励（就如同教育市场上人们有受到高报自己的能力的激励一样）。在教育市场上高报自己的能力的人常常被发现，这样他们就不能长久地享受高报能力所带来的收益（高工资）。但是在资本市场上，往往很难确认谎报的情况，而且等到确认的时候，证券的卖家早已携款潜逃了。因此资本市场就比教育市场更加有必要引入外部的“公共”机构来验证信息。在资本市场上，这类机构包括投资咨询顾问、股票交易公司等。[22]

---

21 当然，别的人可能比企业家拥有更多的关于企业前景的信息。但是那样的话，这些人会接管企业然后披露企业信息。然而倘若企业最初的所有者知道有人有兼并动机，他们是不会愿意出售的。

22 之所以要区分直接生产的信息和认证的信息，主要依赖生产不同的信息有不同的成本以及有些信息可能与好几个企业都相关。通用公司可能拥有提供通用公司生产特征信息的比较优势，而证券公司能够提供整个汽车市场的需求。

5. 等级甄别信息的一个例子

假设每个人都是风险规避的，而且以资产组合回报的期望和方差来评估资产。还假设证券的种类足够多，可以完全地将风险分散开来。假设有两种证券，第一种类型的平均收益为 θ_1，第二种类型的平均收益为 θ_2，$\theta_1 > \theta_2$。它们的方差相同。而且每一种证券的数量相等。为了简单起见，假设所有证券都是独立同分布的。并且其收益是服从正态分布的。所有人的效用函数都设为 $-e^{-\alpha Y}$的形式（绝对风险是常数）其中 Y 是收入。因此，个人追求最大化：

$$\bar{Y} - \frac{\alpha\sigma_y^2}{2} \tag{30}$$

其中 $\bar{Y}$ 是收入均值，σ_y^2 是收入方差。最初每个人都拥有一家企业的股份。他仅仅只是知道这个企业的类型，因此，只有他能够向市场提供这个企业的信息。而提供信息的成本为 c，假设 c 既不是太大也不是太小，[23]

$$\left(1 - \frac{\theta_1 - \theta_2}{2\alpha\sigma_y^2}\right)(\theta_1 - \theta_2) > c > (\theta_1 - \bar{\theta})\left(1 - \frac{\theta_1 - \bar{\theta}}{2\alpha\sigma_y^2}\right) \tag{31}$$

这里一共有两个均衡。第一个是没有信息披露的均衡，所有证券的价格都是一样的，个人将保留一部分其所拥有企业的股份 λ，然后把剩余的资产组合分散到所有其余的企业中。λ 的值是下面的问题的解：

$$\max \bar{\theta} + \lambda(\theta_1 - \bar{\theta}) - \frac{\alpha\sigma^2\lambda^2}{2} \tag{32}$$

其中 $\bar{\theta}$ 是市场里证券的平均收益。此时，我们假设低产量的企业主与高产量的企业主持有相同的内部股份。[24]这就意味着随机选择的市场股份的平均收益是$\frac{\theta_1 + \theta_2}{2}$，倘若市场内企业数目多，那么就可以忽略企业行为对市场平均收益的影响 。因此，有

$$\lambda^* = \frac{\theta_1 - \bar{\theta}}{\alpha\sigma^2} \tag{33}$$

23 这些不等式的作用很快就能够显示出来。

24 此假设的合理性将在下一部分里证明。其基本的论证是很简单的：如果企业最初所有者持股份额是可以观察的变量，而且如果低生产能力的企业持股份额小于 λ，那么他们就向市场发出了低生产能力的信号，这样就会使市场价值大受影响。假设更小的低生产能力的企业持股份额不会影响这个结果。

由式（33）可以很容易地得到拥有第一种类型企业的企业家的效用是

$$\bar{Y} - \left(\frac{\alpha}{2}\right)(\sigma_y^2) = \bar{\theta} + \frac{1}{2}\left[\frac{(\theta_1 - \bar{\theta})^2}{\alpha\sigma^2}\right] \tag{34}$$

他的福利略高于普通人。而高的程度取决于他的风险偏好 α 以及他的内部股份收益的方差。如果他采取甄别信息的策略，那么他的资产会上升 $\theta_1 - \bar{\theta}$，而他就可以将资产风险完全地分散掉。因此有：

$$\bar{Y} - \frac{\alpha}{2}\sigma_y^2 \approx \theta_1 \tag{35}$$

而在式（31）的假设条件下，他这样做是无利可图的。

第二种均衡是信息完全披露的均衡。两种证券的相对价格为 θ_1/θ_2，对于高收益的企业，其收入扣除甄别信息的成本后是

$$\tilde{Y} - \frac{\alpha\sigma_y^2}{2} \approx \theta_1 - c \tag{36}$$

而低收益的企业，是：

$$\bar{Y} - \frac{\alpha\sigma_y^2}{2} \approx \theta_2 \tag{37}$$

如果他不采取甄别信息的策略，那么他期望效用是

$$-\exp - \alpha\left[\theta_2 + \frac{1}{2}\frac{(\theta_1 - \theta_2)^2}{\alpha\sigma^2}\right] \tag{38}$$

因此甄别信息是有利可图的。

这个例子阐明了信息不完美市场的几个重要特性：1）可能存在多重均衡；2）其中某些均衡是帕累托占优的，在这个例子里，可以证明甄别信息均衡要帕累托优于不甄别信息均衡；3）第二种类型证券的持有人对第一种类型证券的持有人产生了外部性。倘若所有证券都是优质的（第一种类型），那么第一种类型证券的持有人的期望效用就是 $-\exp(-\theta_1\alpha)$，在两种均衡里，第一种类型证券的持有人的福利都降低了。另一方面，相对于如果所有证券都是第二种类型的情况，第一种类型证券的持有人的损失要高于第二种类型证券的持有人的收益。这正是甄别均衡模型的特征（在完全信息条件下），实际上第二种类型证券的持有人的福利状况要低于不存在第一种类型证券的持有人的情况，而后者也很不情愿地遭受了福利损失：所以这种外部性完全是耗散性的。然而，即使相比不甄别信息均衡，第一种类型证券的持有人的“等收益损失”是：

$$(\theta_1 - \bar{\theta})\left(1 - \frac{1}{2}\frac{(\theta_1 - \bar{\theta})}{\alpha\sigma^2}\right) \tag{39}$$

而第二种类型证券的持有人的收益仅仅是

$$(\bar{\theta}-\theta_2)-\frac{(\theta_1-\bar{\theta})^2}{\alpha\sigma^2} \tag{40}$$

因此净损失为

$$\frac{(\theta_1-\theta_2)^2}{4\alpha\sigma^2} \tag{41}$$

6. 匹配甄别的一个例子

上文详细论述的例子中包含的一类证券在客观上要好于另一类。而在将要讨论的例子里，不同的人对于将来可能发生的自然状态有着不同的看法，这自然会对基于事前的期望效用最优化所选择的股票产生影响。有这样一个企业，他的生产技术要么是 Y 要么是 Z。简单起见，假设个人都是风险中性的，但是不能卖空证券。[25]人群里共有三类人：A 类型的人认为：如果该企业采用 Y 的技术，就会有很高的利润，如果该企业采用 Z 的技术，利润就会很低；B 类型的人认为恰恰相反：Z 技术的期望利润要高，而 Y 技术的期望利润要低；C 类型的人认为两者的期望利润是相等的。在缺乏具体技术信息的情况下，人们认为使用两种技术的可能性相等。准确地说，令 $\prod^A(Y)$ 为 A 类人认为 Y 技术被采用时的期望利润，而 ω^A 为 A 类人的总财富水平。假设另外还存在一个投资机会（证券 Q），每个人都认为其期望利润为 γ^*，于是假设：

$$\prod^A(Y)>\prod^C(Y)>\prod^B(Y),\omega^A<\frac{\prod^C(Y)}{\gamma^*};$$

$$\prod^B(Z)>\prod^C(Z)>\prod^A(Z),\omega^B<\frac{\prod^C(Z)}{\gamma^*};$$

$$\frac{\prod^A(Y)+\prod^A(Z)}{2}<\prod^C(Y)=\prod^C(Z),\omega^C>\frac{\prod^C(Y)}{\gamma^*};$$

$$\frac{\prod^B(Y)+\prod^B(Z)}{2}<\prod^C(Y)=\prod^C(Z) \tag{42}$$

因此，在缺乏信息的条件下，相对于拥有任何企业，A 类人和 B 类人都更喜欢证券 Q。于是只有 C 类人才会去拥有企业。按照假设条件，他们的财富要高于该企业的价值（该企业的价值为$\frac{\prod^C(Y)}{\gamma^*}$，这要比 C 类人的总财富

25　这些假设是很容易修改的。

值小)。C 类人把剩余的财富，以及 A 类人和 B 类人把所有的财富都投资于证券 Q，这样每一类人的期望收入（EI）就是

$$\mathrm{EI}^A = \gamma^* \omega^A$$
$$\mathrm{EI}^B = \gamma^* \omega^B$$
$$\mathrm{EI}^C = \gamma^* \omega^C \tag{43}$$

现在假设信息是完美的，比如，每个人都知道 γ 技术被采用了。这样 A 类人就会全部购买该企业的证券。但是，由假设可知，他们的财富不足以购买全部的证券，因此 C 类人就成为了那个“边际买家”。于是，该企业的价值依然是 $\frac{\prod^C(Y)}{\gamma^*}$。而期望收入由下式给出：

$$\mathrm{EI}^A = \frac{\prod^A(Y)}{\prod^C(Y)}\gamma^* \omega^A > \gamma^* \omega^A$$
$$\mathrm{EI}^B = \gamma^* \omega^B$$
$$\mathrm{EI}^C = \gamma^* \omega^C \tag{44}$$

因此就存在着帕累托改进的可能性：A 类人的收入增加了而其他类的人也没有损失。(在期望效用的意义上是这样)。如果信息的成本小于

$$\left(\frac{\prod^A(Y)}{\prod^C(Y)} - 1\right)\gamma^* \omega^A, \tag{45}$$

那么获取信息就显然是值得的。但是要注意尽管信息是有价值的，却绝不会对企业的价值产生影响。因此，企业就没有披露信息的激励——尽管这样做几乎是没有成本的。A 类人存在着联合起来组成“消费者共同体”的激励来获取信息；但是由于他们很难互相确认出彼此的身份，所以不大可能成功(因此，尽管存在着一个天然的“厂商”共同体来为购买证券提供信息，即企业本身，但是却不存在一个相应的天然消费者共同体)。就算 A 类人能够互相确认出彼此的身份，但是由于存在着经典的公共品难题（搭便车），使得对该信息的投资很可能是不足的。

八、自选择的均衡

前面部分我们提出给定企业的生产能力，高生产能力企业的企业家受到向市场提供本企业生产能力信息的激励，然而，这里存在着置信力这一基本问题——企业提供的信息（或从企业购买来的）并不是很可信的。众所周知，行动比言语更有说服力。在近期的甄别均衡的文献里，传递了信息的行为（例如，企业的生产能力）被称为自选择机制（Stiglitz，1982）愿意增加

自己在本企业的持股比例的企业家可能传递了这样的信息：该企业的市值被低估了。当然，如果投资者相信这些的话，那么低生产能力的企业就倾向于模仿高生产能力的企业的行为。那么企业家所持有的本企业的股份额也就无法传递信息了。必须说明的是，自选择机制的分析本身是个很复杂的问题。实际上，我们将会看见，甚至很难建立一个适当的均衡定义。自选择机制的关键性质是在于不同人群的相对成本曲线（无差异曲线）是不同的。增加本企业的持股比例的企业家必须承担没有分散风险的成本（风险增加了）。如果证券的市值被低估了，就存在着获利的机会——增加持股比例会提高平均收入。如此一来，高生产能力的企业所承担的成本就要小于低生产能力的企业，正是这个差别使得在某些情况下企业家所持有的本企业的股份额能够实现自选择机制。

为了更加正式地证明这一点，我们回到第七部分第5点的那个简单的例子。令 p_i 为第 i 种企业家出售该企业的价格，将市场上股票的平均利润单位化为1，图1在（λ，p）空间中描述了消费者的无差异曲线（λ 是持有股票

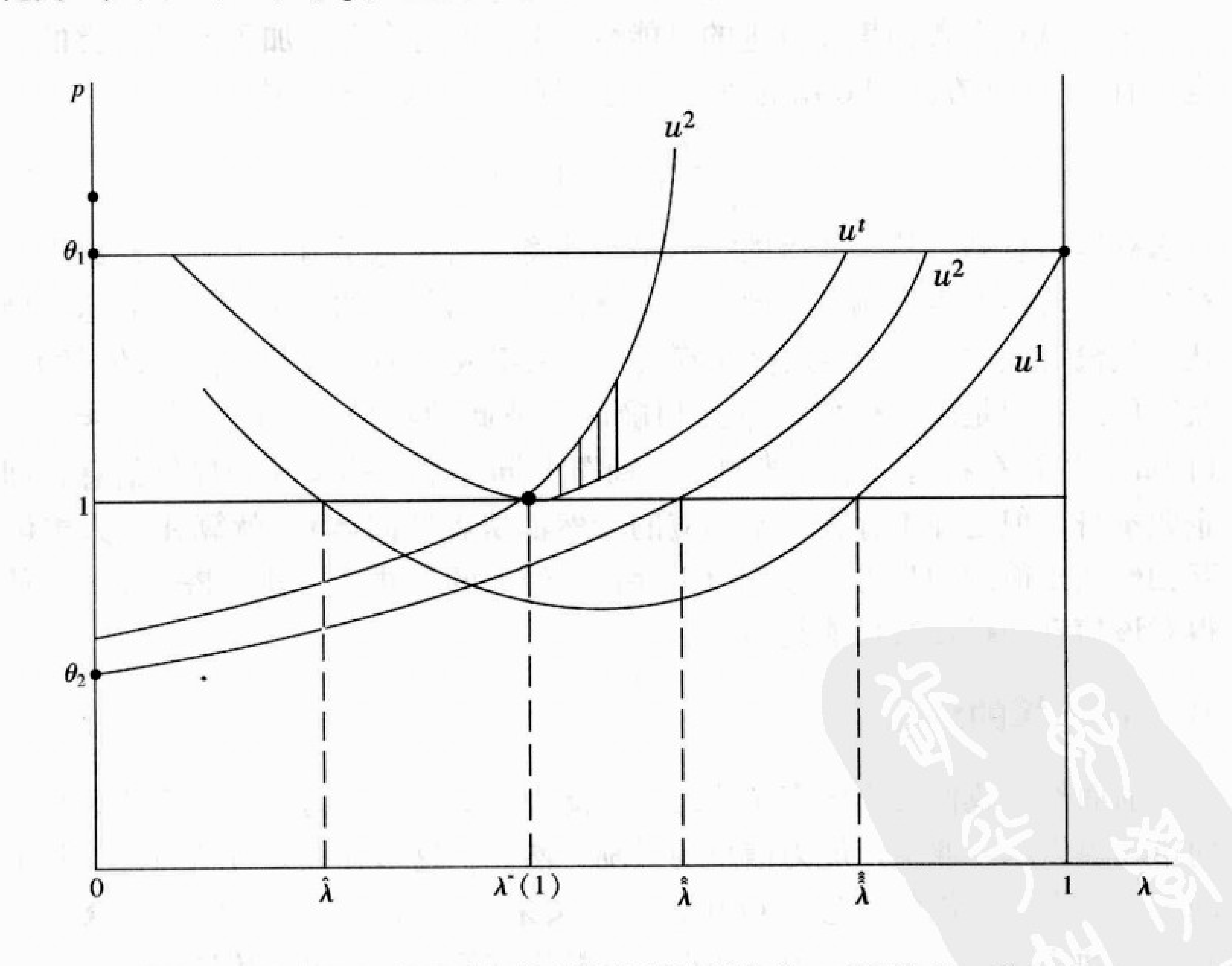


图1 企业作为一个甄别机构所持有的股份比率：对风险的一致态度

份额），显然，当 $\lambda \gtrless \lambda^*(p)$ 时有

$$\left(\frac{dp}{d\lambda}\right)_u \gtrless 0 \tag{46}$$

其中 λ^* （p）是最优的股票持有份额（可以参见式（33））。并且还有下式成立：

$$\left(\frac{dp}{d\lambda}\right)_{\bar{u}^2} > \left(\frac{dp}{d\lambda}\right)_{\bar{u}^1} \tag{47}$$

低生产能力企业的企业家倘若要增持内部股份就必须使其股票的价格上涨更多以补偿他的风险。

1. 混同均衡（Pooling Equilibrium）

首先假设 λ 是可以观察的。考虑没有信息披露的情况，当 $\lambda_1 = \lambda_2$ 时（Rothschild and Stiglitz，1976，将其称为混同均衡），显然，均衡中有 $p=1$ 而且 λ 的最优值正是 λ_1（在高生产能力的企业看来）。

然而，这是一个均衡吗？显然，好企业是愿意发送信号表明自己是优质的。假如有个优质企业声称愿意以高于 1 的价格 p 出售 $\hat{\lambda}$ 份额的股票，（$\hat{\lambda}$，p）位于图 1 的阴影部分——高于第一种类型证券的持有人的无差异曲线，低于第二种类型证券的持有人的无差异曲线。投资者可能倾向于觉得只有第一种类型企业家——高生产能力企业的所有者——才会愿意这样做，因此，这个企业一定是生产能力高的，它的市值实际上是 θ_1。但是投资者同样可以推测说，只要企业家（两种类型都是）相信投资者觉得只要是不愿意出售 $\hat{\lambda}$ 份额的股票的企业实际上都是生产能力低的，那么他一定会持有很大份额的内部股份，并且以很高的价格出售剩余股份。在这种情况下，投资者会认为企业家声称愿意持有大于 $\hat{\lambda}$ 份额的股票是不能传递有效信息的。

到底是何种均衡似乎是由投资者对于两类企业非均衡行为的预期决定的。例如，假设投资者相信只有第二种类型企业会偏离 $\lambda = \lambda^*$。那么显然第一种类型企业就永远不会偏离均衡，因此这个均衡是自我实现的。而且如同我们在图 1 中所画出第一种类型企业家的无差异曲线与纵轴的交点在 θ_2 之上那样，第二种类型证券的持有人偏离均衡的行为也是无利可图的。于是，（λ^*，1）构成一个均衡，并且也和预期一致——投资者永远不会遇见与其预期相冲突的行为。这种均衡的定义的一个难题在于可能会存在很多的均衡。λ 的最优值的确定可能会是非常主观的（也许符合现实情况），例如，图 1 中画出了第一种类型证券的持有人的无差异曲线经过点（1，θ_1），与 $p=1$ 线相交于 $\hat{\lambda}$ 和 $\mathring{\lambda}$。在 $\hat{\lambda}$ 和 $\mathring{\lambda}$ 之间的任何一点都可以是均衡，只要预

期给定如上（$\overset{*}{\lambda}$ 是经过点（0，θ_2）的第二种类型证券的持有人的无差异曲线）。

多重均衡存在的事实并不是说任何情况下都是均衡。比如假设类型是在 θ_2 和 θ_1 之间的连续分布，而投资者觉得“差”的企业，那些 $\theta_2<\theta<\hat{\theta}<\theta_1$ 的企业会通过提高自身持股份额（超过 λ^*）来掩饰自己是差企业。因此，他们相信偏离均衡的平均收益是$\frac{\hat{\theta}+\theta_2}{2}$。这里假设企业服从均匀分布。但是很明显，这些预期不是一致的：区间里最好的企业是不会偏离的，所以偏离均衡的平均收益是小于$\frac{\hat{\theta}+\theta_2}{2}$。在这种情况下，唯一的一致预期就是只有最差的企业会偏离均衡。

但是，如果允许好企业的企业主拥有不同的风险偏好，那么均衡时就会观察到一组（λ，p），而且如果风险偏好的差异足够大的话，均衡时就有可能观察到所有的 λ 值。(在 0 和 1 之间）我们还可以讨论一下理性预期（rational expectation）均衡的意义，这个概念要比前面所述一致预期概念更强。

这样的均衡是很容易描述的：有 p（λ）函数，只要给定价格作为自己持有股份份额的函数，满足企业最大化其效用的性质；若给定 p（λ）函数（见图2)，选择 λ 的企业产出的价值就等于 p（λ）。在这里所讨论的简单模型里，低生产能力企业的企业家都具有相同的风险偏好，p（λ）曲线和其中一个第二种类型证券的持有人的无差异曲线完全一致。假定 p（λ）曲线要比第二种类型证券的持有人的无差异曲线更陡（见图 3)。那么所有的第二种类型证券的持有人都会选择 $\bar{\lambda}$ 使得 $\bar{\lambda}=p^{-1}$（θ_1），因此，当（$\lambda<\bar{\lambda}$）时，p（λ）就必须等于 θ_1；当 $\lambda=\bar{\lambda}$ 时，p（λ）就必须小于1。于是，原来的 p 函数就不是均衡的了。

给定 p（λ）函数，第一种类型证券的持有人依照他们各自的风险偏好沿这条曲线被区分开来，他们在曲线上的位置比例正好使得在每一个 λ 值处的平均产出价值等于 p（λ）。

当第二种类型证券的持有人不都具有相同的风险偏好时，构建均衡的 p（λ）函数只是稍稍复杂一点。他代表了第二种类型证券的持有人的无差异曲线集的外包络线。在图 4 中我们阐释了两种偏好情况下的解。风险厌恶第二种类型证券的持有人中选择低的 λ 人越多，那么他们中选择的高 λ 人就越少。第一种类型证券的持有人沿这条 p（λ）的轨迹（圆齿形）被区分开

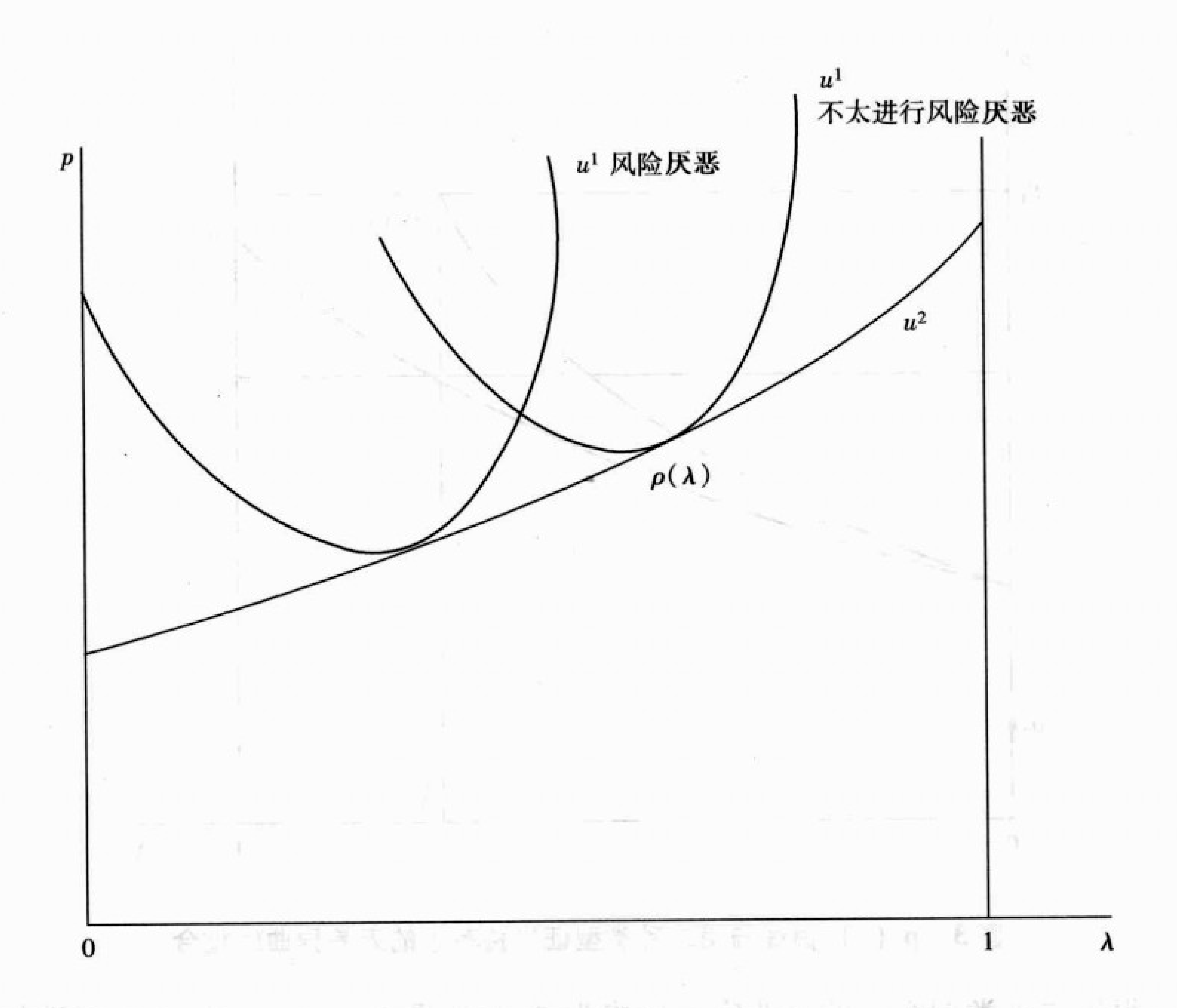


图 2　企业作为甄别机构所持有的股份比率：对风险的不同态度

来（注意即使第一种类型证券的持有人的偏好分布不存在聚点，当 p（λ）不是可微的情况下仍然会产生聚点）。

2. 分离均衡（Separating Equilibrium）

截至目前，我们的分析只是限于混同均衡的情况（那里投资者无法根据企业的行为来区分出企业的类型）以及部分的分离均衡（只能部分地区分出企业的类型）。即使在这些例子下，企业的行为可以传递信息也是具有重要意义的：低生产能力企业的企业家被迫持有更多的内部股份，以使信息不被泄露。混同均衡与以前部分的没有信息披露的情况相对应。

然而可能存在着不同类型的企业采取不同行为的均衡，这样投资者可以依据企业的行为来推断企业的类型。我们把这种情况称为分离均衡。图 5 中给出了一个这样的例子。其中第二种类型企业的企业家将其资产完全地分散了（令 $\lambda_2=0$）。而第一种类型企业的企业家将 λ 定为 $\bar{\lambda}$。$\bar{\lambda}$ 的值足够大，

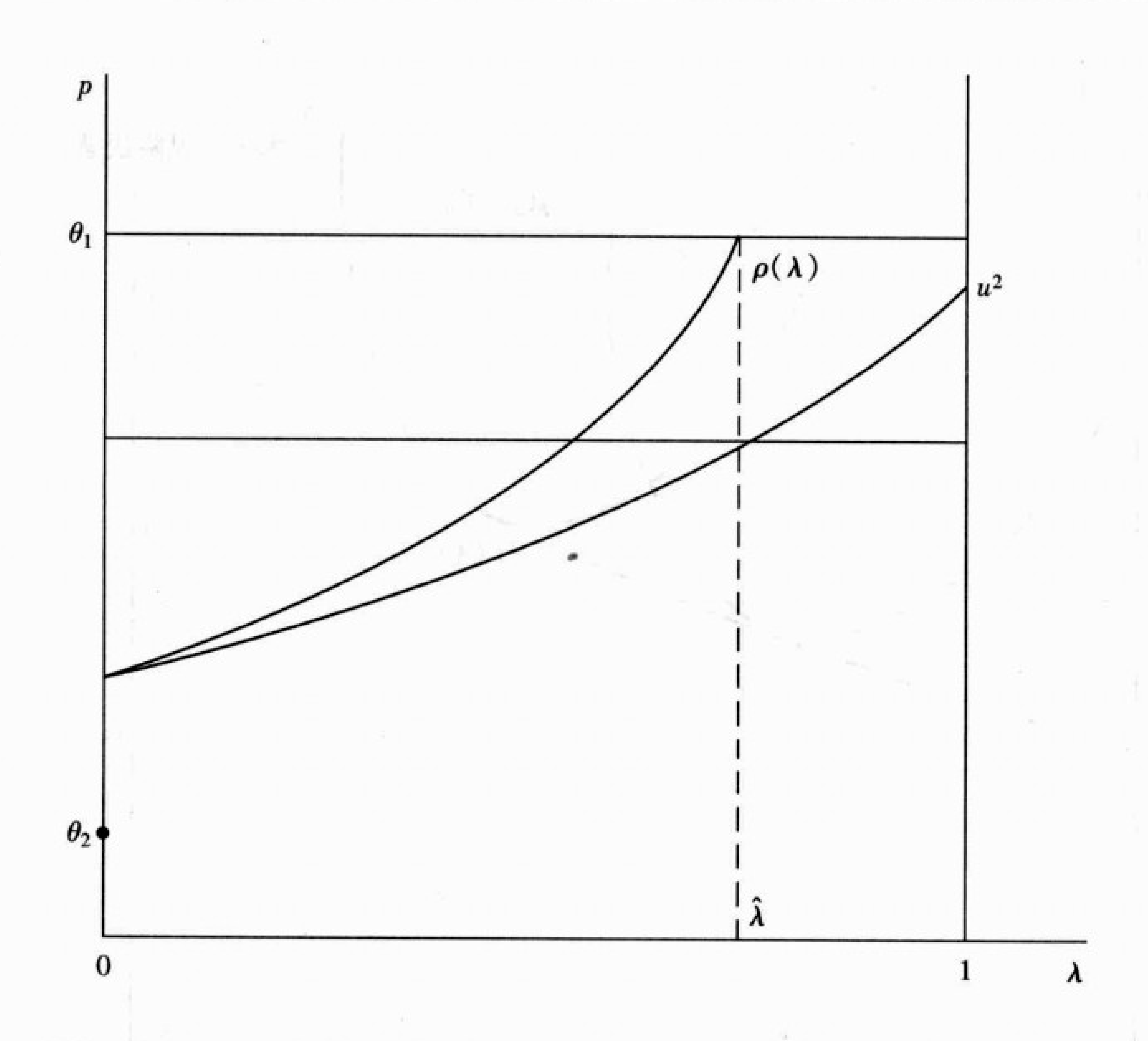


图3 p（λ）曲线与第二种类型证券持有人的无差异曲线重合

使得第二种类型企业的企业家在价格为 θ_2 时选择 $\lambda_2=0$，价格为 θ_1 时选择 $\lambda=\bar{\lambda}$。任何大于 $\bar{\lambda}$ 的 λ 值（通过（θ_2，0）第二种类型企业家的无差异曲线与$p=0$交点）都会产生分离均衡（自我选择）。

同样地，是否存在均衡取决于关于投资者预期的特定假设。如果假设他们认为偏离均衡 λ_1 值的只可能是第二种类型企业，那么显然任何满足 $\lambda\geqslant\bar{\lambda}$ 的 λ 都可以构成均衡。而且，这里的预期和前面的定义是一致的——它们永远不会被第一种类型企业的行为推翻。另一方面投资者同样可以认为大的 λ 值更加可能是与第一种类型企业相关联的，而小的 λ 值是与第二种类型企业相关联的，选择任何一种（λ，p）值都会改变均衡时福利函数的位置。例如当 λ 的值为 $\lambda_2>\lambda>\bar{\lambda}$ 时，如果第二种类型企业从点（0，θ_2）移动到点（λ，θ_1）就会有严格的福利损失，而第一种类型企业从点（λ_1，θ_1）移动到点（λ，θ_1）就会使福利严格增加。投资者可以推断所观察到的在点（λ，θ_1）的企业实际上是第一种类型企业。但是这样（λ_1，θ_1）就不是一个均衡了。因为所有的第一种类型企业在知道这个信息后都会降低 λ 的值。

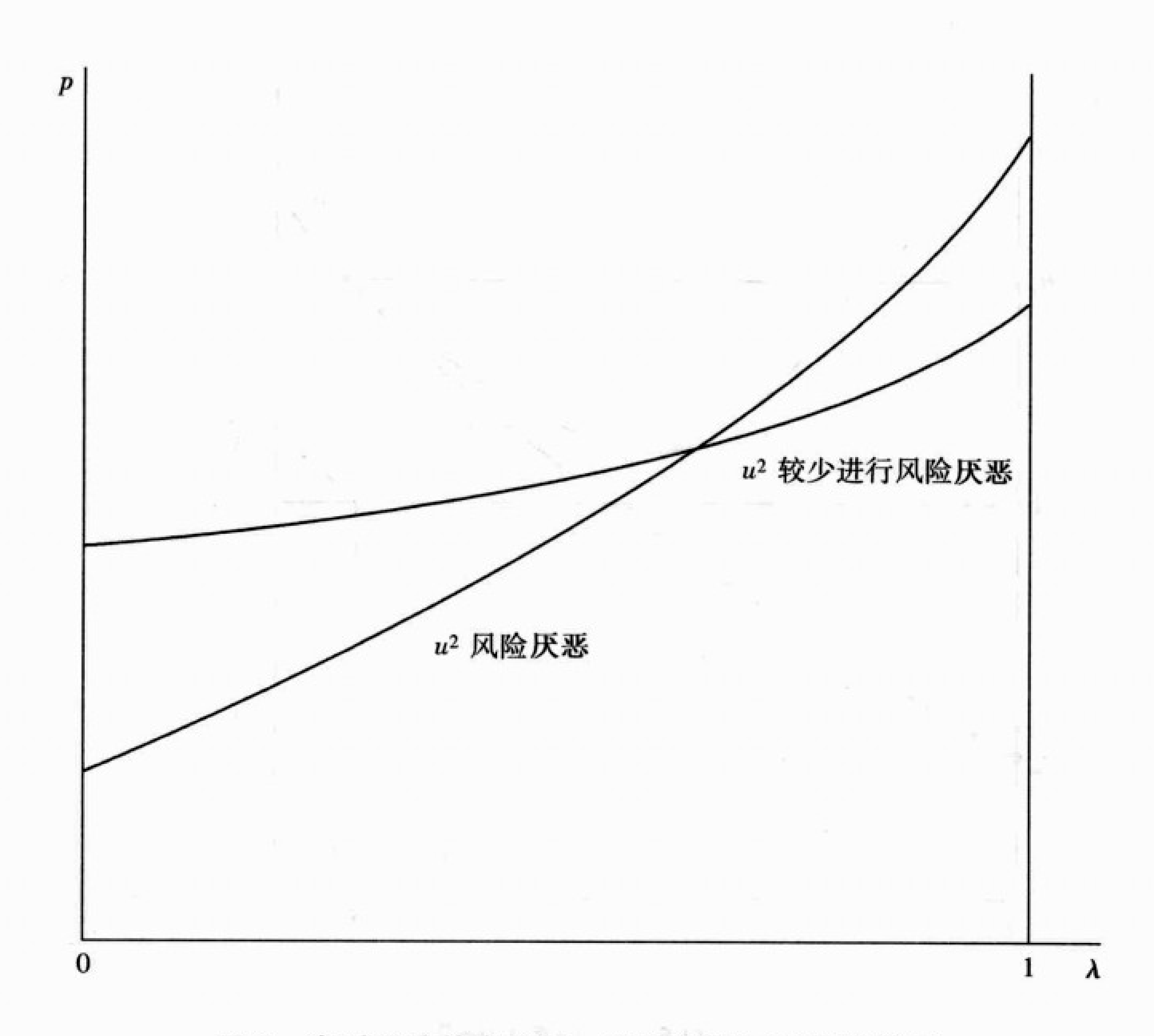


图 4　当对风险的态度不一致时价格与留存率关系

另一方面，（0，θ_2）和（$\bar{\lambda}$，θ_1）可能构成一个均衡。把 λ 的值从 $\bar{\lambda}$ 降低一点点会使得所有人的效用都升高。考虑一种简单的期望函数形式：它假定，如果存在着任何可以增加所有人效用的行动，那么所有人都会加入的，于是行动后唯一的均衡价格就只能是 $p=1$。显然，在这样的预期里，图 6 的分离点不再构成均衡。因为在点（λ^*，1）所有人的效用都比分离点要高。然而，在这样的预期里，图 5 的分离点却能够构成均衡，因为第一种类型企业的企业家的无差异曲线通过点（$\bar{\lambda}$，θ_1）而且位于 $p=1$ 线之上。

3. 小结

这一节，我们已经阐明了自持股份份额是怎样传递企业生产能力的信息的，以及市场了解到这一点是如何影响企业行为的。我们证明了可能存在着多种均衡：企业行为可以完全地传递企业的特征信息；也可以部分地传递企业特征信息；还可以不传递企业特征信息。多重均衡的关键问题在于投资者是如何形成预期的假设。在很多情况下，重要的是如何推断均衡下企业的行为。这里就存在着许多可能的均衡，它们都与预期是一致的。在其中的一个

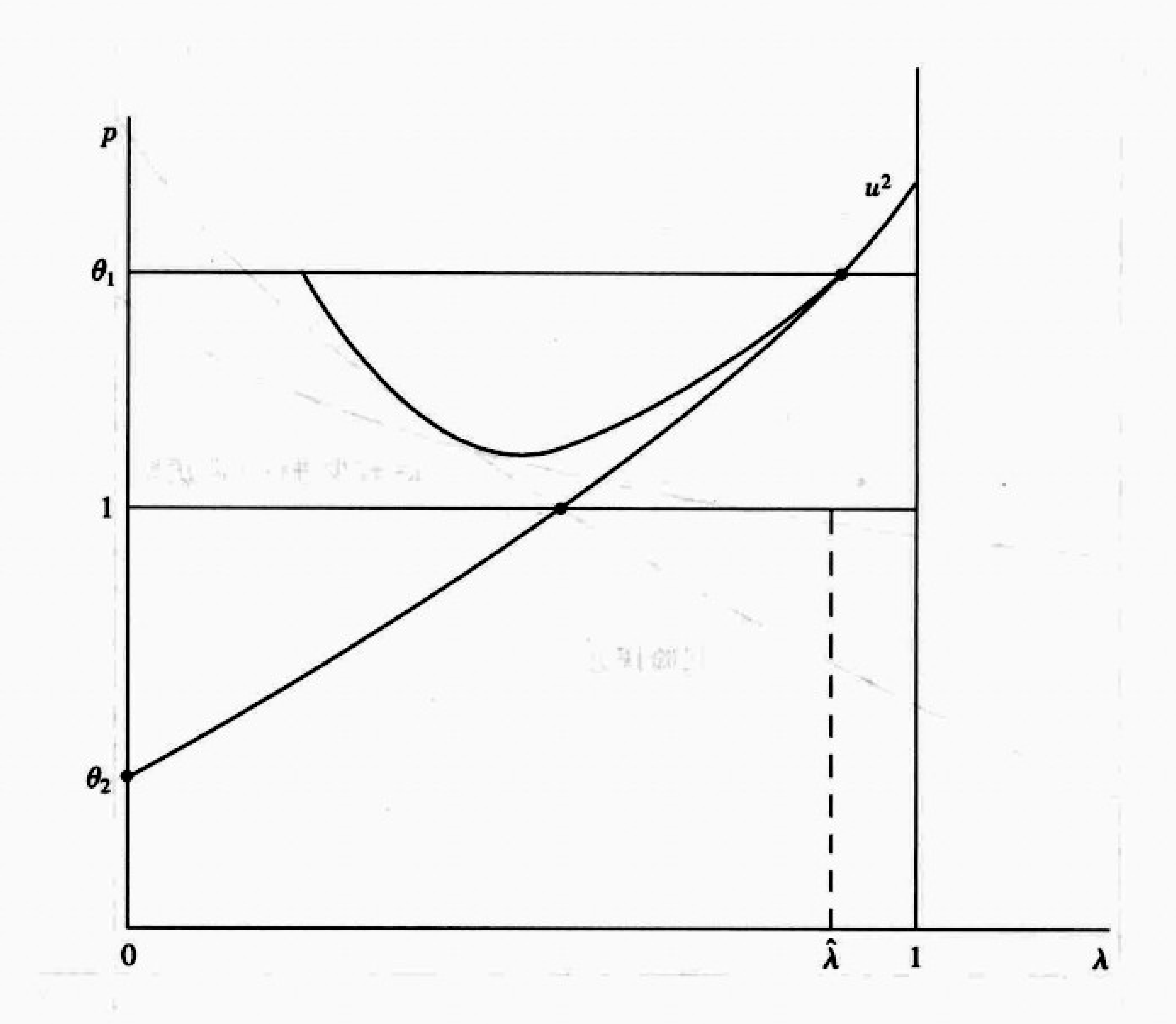


图 5 "分离均衡"

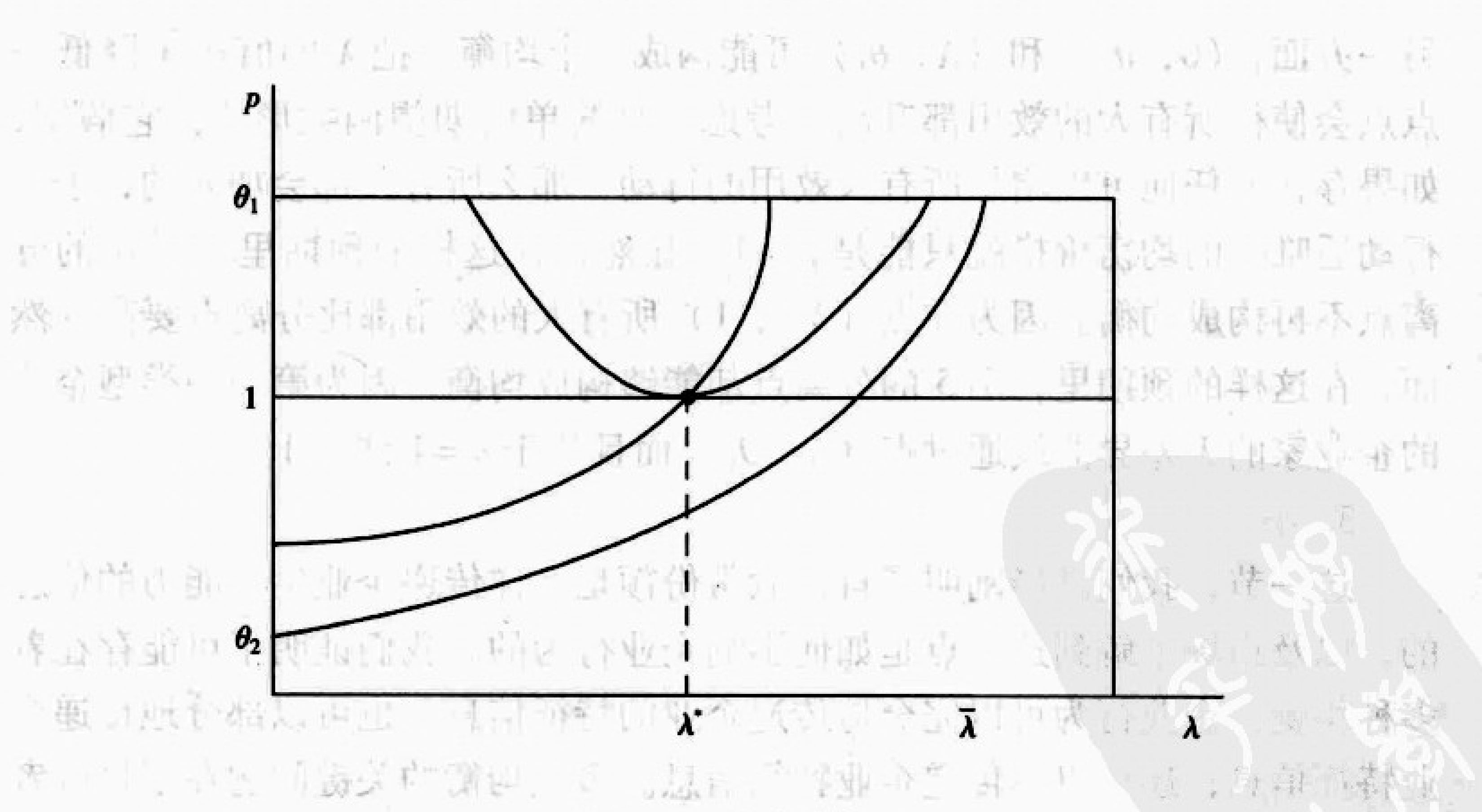


图 6 无分离均衡

情况下，我们可以找到完全理性预期的均衡。

然而，在所有讨论的情况下，高生产能力企业的企业家（在混同均衡以及部分混同均衡中的所有企业）都持有相当大比例的内部股份。这一点对于理解这些企业的行为有着重要意义，下面我们就会看到。

九、不完全分散风险的企业投资行为的意义

前面我们曾经证明，尽管个人能够完全地分散风险，但是由于信息差别他可能不会这么做。注意到尽管个人对于缺乏信息的证券是风险中性的，但是他有可能对他了解到的几家企业的股票显示出很强的风险厌恶倾向。这是具有重要意义的。就算市场上企业的市值不完全取决于其平均利润（企业规模），企业（人们很可能会想，企业是否被有信息的人所持有的）依然会显示出风险厌恶倾向。一般而言，企业并不会最大化其净市值。[26]比如，假设企业可以通过增加投入来提高产出 I，而生产函数的风险是满足可乘性的：$Q=g(I)\theta$，$g'>0$，$g''<0$。g 是可观察的，但是它的分布依赖于企业的类型。而这个信息是外部的投资者所不知道的。市场将该企业看作平均水平的企业，其市场价值就是 $V=\theta[g(I)-I]$[27]。于是，第六部分例子中的企业家知道 $\theta=\theta_1>\bar{\theta}$，他选择 I 以满足下式：

$$\max[\lambda\theta_1+(1-\lambda)\bar{\theta}]g(I)-\bar{\theta}I-\frac{\alpha}{2}\lambda^2\sigma^2(g(I)-I)^2 \tag{48}$$

假设有足够多的企业可以令他完全地分散风险，即，

$$[\lambda\theta_1+(1-\lambda)\bar{\theta}]g'-\bar{\theta}-\alpha\lambda^2\sigma^2(g-I)(g'-1)=0 \tag{49}$$

或者，

$$(g'-1)=\frac{\lambda(\theta_1-\bar{\theta})}{\alpha\lambda^2\sigma^2(g-1)-(\lambda\theta_1+(1-\lambda)\bar{\theta})} \tag{50}$$

二阶条件保证了分母为正。而企业价值最大化要求

$$g'=1 \tag{51}$$

因此，企业的产量是在市值最大化水平之下的。

企业的市场价值的确是企业投资决策的一个考虑因素，但是由于它不能反映出作为投资项目的企业的实际生产能力，因而仅仅只是一个决定因

---

26 如果企业最初所有者出售的股份份额并不是公共信息，那么平均水平以下的企业就会显示出风险厌恶的倾向。结果就是收益方差大的企业的平均收益不会超过甚至可能小于方差小的企业的平均收益，这样可能会使得人们在权衡真实的收益方差做出错误的推测。

27 这个公式实际上假设了企业会借钱为 I 融资，而支付平均资本回报 $\bar{\theta}$。

素而已。由于经理兼企业主很可能比市场多知道一些关于作为投资项目的企业的实际生产能力，因此，他们不会仅仅依靠市场的判断来决策的。实际上，式（50）已经阐明了投资决策的关键在于企业家与市场之间的信息不对称。

注意到在前面的论述中我们已经证明了 λ（原始企业所有者所持有的股份比例）是如何传递信息的，那么 I（企业的投资数额）也可以传递企业家对于本企业生产能力的看法的信息。因此，它具有重要的经济含义：如果投资规模可以传递信息的话，那么要么低生产能力的企业会过度投资，来掩盖它们与高生产能力的企业之间的差别（在不甄别均衡中），要么高生产能力的企业会过度投资，显示它们与低生产能力的企业之间存在差别。[28]

不完全的分散风险源于不完全信息，这一点不仅对投资水平具有重要意义，而且对于企业技术的选择同样重要（这往往不太容易观察）。企业不仅关注自己的风险（原则上不会影响其市场价值）而且还关注本企业与整个市场的联系（会影响其市场价值）。因此，相比隐含完全甄别的传统模型，这个风险承担模型所预测的企业行为与现实更加接近。

福利含义

前面我们曾经提到一个普遍接受的观点，就是在纯粹的交易经济中，投资信息是一种社会损失——它只能够让人们损人利己 。第七部分的分析为此观点提供了一定的支持，那里我们构造了一个例子，存在着没有信息披露的均衡帕累托优于完全信息披露的均衡。然而就算是在纯粹的交易经济中，我们也证明了信息不完美是有成本的——个人投资风险可能不会完全分散。相比信息完美的情况，信息不完美导致了个人被迫承担更多的风险。但是一旦引入生产，信息不完美引起了更多的损失。例如我们证明了投资风险不能完全分散会影响投资水平以及技术的选择。实际上，第七部分的例子只要修改一下就可以用来证明当面临如此的生产选择时，依然存在着多重均衡。不过，完全的信息披露均衡可能会比没有信息披露均衡更具有帕累托最优。

十、新企业

前面我们假设了企业的数量是一定的（尽管种类繁多）。但是资本市场

28 这样模型就阐释了规模最大化假设的基本原理：投资的水平可以作为一种信息甄别机制。

的一个重要的功能就是为新的企业提供资助。"牛皮"的供给函数具有无限的弹性，往往投资者还必须识别哪些是好的投资机会，哪些仅仅是吹牛而已。事实上，要是真有人随机地把自己的存款投资于新发行的股票，那么他的收益很可能是零（如果市场知道他的做法）。尤其是他绝对比不上那些事前研究过市场的人。

关于这一点有两种反对意见。第一，如果没有信息的人仅仅模仿拥有信息的人，难道他们就不能做到至少和拥有信息的人一样吗，何况他们还不用付信息成本？不尽然。只要信息的传递具有时滞（例如，只要股票交易人在下单交易之前不透露拥有信息的人的行动）。当然拥有信息的人购买新企业股票的行为确实传递了信息。正如我们前面指出的那样，他会使得下一期该股票的价格上涨（当信息披露后）。正是由于高的利润使得拥有信息的人愿意进行信息投资。

而前面论述中假设企业以固定的价格有弹性地提供证券，这会产生严重的问题。任何随机地购买证券的人最后所得到的大部分都是吹牛公司的股票，显然要比拥有信息的人差很远。实际上，企业一般会提供一个固定的证券数量。假如企业愿意出售全部的证券，那么任何随机地购买证券的人都能够获得证券市场的平均收益。那么投资于企业信息的人就不如没有信息的人（由假设证券的供给是没有弹性的，与先前分析的纯粹交易经济的情况等价）。

这个结果很令人不快，要是拥有信息的人投资了信息却无法得到回报，那么他们就会丧失获取信息的激励；市场中会充斥着"牛皮"泡沫，产权投资市场就会失去其功能。大多数新股是以固定价格发行固定的数额，如果存在超额需求就实行配给。如果配给的数额与需求相当，那么拥有信息的人的回报就能够胜过缺乏信息人的回报。倘若"牛皮"证券的供给是无限弹性的，而且缺乏信息的人无法同时观察到拥有信息的人的需求，那么他们就不会购买新股。

这个论点解释了为何证券产业往往乐意以低于市场出清价格发行新股，但是没有解释为何任何一个企业都会乐意以低于市场出清价格发行新股。平均而言，拥有信息的投资者一定能够从投资信息中获得回报；但是一旦他发现任何证券的价格低于市场出清价格，他就会购买。这样，每个企业都应该提高新股发行价格至市场出清价格（给定拥有信息的投资者数目）。但是如此一来投资者就不会有获取信息的激励了。

如果企业不能准确地预测出均衡价格（市场出清价格），那么它们有可

能制定一个足够低的新股发行价格，使得很多情况下都会产生配给（*rationing*）。[29]在配给条件下，证券的价格在初次交易后会上涨，拥有信息的投资者就能够从投资信息中获得回报。[30]

当然，新股发行还有别的市场策划方式（比如采取拍卖的方式）。为何以固定价格配给新股的方式在证券市场占有统治地位还有待进一步研究。

十一、结论

本文尝试建立一个关于信息与资本市场的一般经济理论。下面是本文得到的一些重要结论：

1. 在由理性的人组成的纯粹赌博市场上，是不存在交易均衡的，即便在人们拥有不同信息的条件下也是如此。

2. 尽管纯粹的证券交易市场不是纯粹赌博市场，由于人们禀赋中的不可贸易资产（人力资本）与市场上的不同证券有着不同的相关性，市场上大多数交易以及获取信息的行为并不是源于人的禀赋；相反，上述交易和行为可能归因于一部分投资者的非理性行为以及其他理性投资者想要利用这种非理性的行为。

3. 获取信息以及披露信息的私人价值与社会价值是有着很大的区别的。在纯粹的（证券）交易市场上，被低估企业的所有者可以通过提供企业证券的信息来获取个人收益。而这种收益完全是收入再分配性质的。在不完美信息条件下，有的企业市值会被低估，这会导致这些企业的所有者不能够完全地分散证券风险。这样，由于信息不完美导致所承担的额外风险就形成了社会成本。

4. 企业的行为——企业家自己所持有的股权比例、资产负债比例、投资水平——都可能传递企业特征的信息，从而影响企业的行为。例如，差企业的企业家不得不持有很多的内部股份，以便信息不被泄露，好企业的企业家被迫持有更多的内部股份，以使信息得以披露。

5. 并未完全将投资组合分散化的事实有更深的含义：企业并不会采取

29 可能是由于新股的发行方并不比某些潜在买家有更多信息，也可能是由于拥有信息的人数非常少，使得获取特定的证券信息的人数和信息量都具有不确定性。倘若投资者是风险规避的或者不能自由地进入市场，就算他们观察到了一种证券被低估，他们都只会购买有限数额的证券。

30 注意到根据第八部分的论述，最初的股东会持有相当大份额的企业股份；于是，他们不仅仅会关注新股发行时的价格，还会关注其后市场的价格。我们还没有分析最优价格（和存在配给时相应的最优持股份额）。

让市场价值最大化的行为，特别地，它们有规避风险的行为，关注自身风险；而传统的理论认为企业最多只会考虑那些与证券市场相关的风险。

6. 不能认定所有市场均衡都是帕累托最优的，在纯粹交易经济的条件下，我们构造了一个存在着多重均衡的例子。而没有信息披露的均衡要帕累托优于完全信息披露的均衡。而在存在着投资和技术选择的经济里完全可以产生相反的例子。

尽管我们给出的模型是高度理想化的，但是我相信，相比忽视信息或者假定信息是无成本的传统的理论模型，我们给出的模型奠定了一个基础，使得我们能够更好地理解企业和资本市场许多方面的行为。

参考文献

Arrow, K. J., "The Role of Securities in the Optimal Allocation of Risk-Bearing," Review of Economic Studies, 31 (1963), 91 - 96.

Aumann, Robert J., "Agreeing to Disagree," *The Annals of Statistics*, 4, no. 6 (1976), 1236 - 39.

Cass, D., and J. E. Stiglitz, "The Structure of Preferences and Returns and Separability in Portfolio Allocation: A Contribution to the Pure Theory of Mutual Funds," *Journal of Economic Theory*, 2 (June 1970).

Debreu, Gerard, *Theory of Value*. New Haven: Yale University Press, 1959.

Green, J., "Information, Efficiency and Equilibrium," mimeo, Harvard University, 1973.

Grossman, S., "On the Efficiency of Competitive Stock Markets Where Traders Have Diverse Information," *Journal of Finance* (1976), 573 - 85.

——, "The Existence of Future Markets, Noisy Rational Expectations and Informational Externalities," *REStud* (1977).

Grossman, S., and O. D. Hart, "Takeover Bids, the Free Rider Problem, and the Theory of the Corporation," *Bell Journal of Economics*, 11 (Spring 1980).

Grossman, S., and J. E. Stiglitz, "Information and Competitive Price Systems," *American Economic Review Papers and Proceedings* (May 1976).

——, "On Value Maximization and Alternative Objectives of the Firm," *Journal of Finance*, 32 (May 1977), 389 - 402.

——, "On the Impossibility of Informationally Efficient Markets," *American Economic Review*, 70 (June 1980a), 393 - 408.

——, "Stockholder Unanimity in Making Production and Financial Decisions," *Quarterly Journal of Economics*, XCIV, no. 3 (May 1980b), 543 - 66.

Hirschleifer, J., "The Private and Social Value of Information and the Reward to Inventive Activity," *American Economic Review*, LXI, no. 4 (September 1971), 561 - 74.

Kreps, D., "A Note on 'Fulfilled Expectations' Equilibrium," *Journal of Economic Theory* (1977).

Leland, H., and P. Pyle, "Informational Asymmetries, Financial Structure and Financial In-

termediation," *Journal of Finance* (May 1977), 371 –87.

Milgrom, P., and N. Stokey, "Information, Trade and Common Knowledge," mimeo, Northwestern University (April 1979).

Radner, R., "Competitive Equilibrium under Uncertainty," *Econometrica*, 36 (1968), 31 –58.

Ross, S., "The Determination of Financial Structure: The Incentive-Signalling Approach," *Bell Journal* (Spring 1977), 23 –40.

Rothschild, M., and J. E. Stiglitz, "Equilibrium in Competitive Insurance Markets," *Quarterly Journal of Economics* (November 1976), 619 –28.

Salop, S., and J. Salop, "Self-Selection and Labor Turnover in the Labor Market," *Quarterly Journal of Economics* (November 1976), 619 –28.

Spence, M., "Job Market Signalling," *Quarterly Journal of Economics* (August 1973), 355 –79.

Stiglitz, J. E., "On the Optimality of the Stock Market Allocation of Investment," paper presented at Far Eastern Meetings of Econometric Society, Tokyo, June 27 –29, 1970, "Perfect and Imperfect Capital Markets," paper presented to Econometric Society Meetings, New Orleans, 1971.

——, "Information and Economic Analysis," in *Current Economic Problems*, Parkin and Nobay, eds. Cambridge, England: Cambridge University Press, (1975a).

——, "The Theory of Screening, Education and the Distribution of Income," *American Economic Review*, 65, no. 3 (June 1975b), 283 –300.

——, *Information and Economic Analysis*, forthcoming in 1982.

Stiglitz, J. E., and A. Weiss, "Credit Rationing in Markets with Imperfect Information, Part II: Constraints as Incentive Devices," Princeton University Econometric Research Program Memorandum No. 268, 1980.

——, "Credit Rationing in Markets with Imperfect Information, Part I," *American Economic Review*, (June 1981), 393 –410.

终止的激励效应：在信贷市场与劳动力市场中的应用*

对于产出低的工人，企业往往是将其解雇而不是降低其工资；对于产量太低的小佃农，地主往往是收回租地，将其逐出，而不是降低其的产出份额；类似地，对于违约者，银行往往是拒绝继续向其贷款，而不是提高其以后需要支付的利息的利息率。现实生活中存在很多这样的带有意外条款的合同（contingency contract），这些合同规定了不好的结果将导致公司—雇员、地主—承租者或者银行—借款人等关系的终止。本文讨论了三个问题：在何种情况下会使用带有意外条款的合同？这种合同的均衡结构是什么样的？在何种情况下均衡的带有意外条款的合同会导致合同关系的终止？我们认为带有意外条款的合同具有某些很好的激励性特征。在上面所列举的发生终止的每一个例子中，终止的威胁会鼓励一些委托人（雇主、地主、银行）愿意看到的行为发生，而避免一些负面激励和分类效应（sorting effects），如果通过改变合同的某些条款来惩罚代理人（雇员、承租者、借款人），则有可能会导致那样的负面激励和分类效应。例如，在劳动力市场上，降低工资会导致工人努力程度的降低或者更多的工人退出（尤其是较好的工人）。类似地，银行不会用增加利率的办法来惩罚违约者，如果那样，它们的利润很可能下降。

在1981年的文章中，我们提出了两条导致借款人所需支付的利率和银行利润之间存在负向关系的原因。第一，提高利率可能会使贷款申请者的平均风险倾向增加（逆向选择效应）。合同的破产条款规定了借款人所要承担的最大风险为抵押品价值，而规定了贷款人所能获得的回报率的上限即为利率。因此，在具有相同期望总收入的项目中，更具有风险性的项目（即为均值保留展型的项目）对借款人有更高的回报率而对贷款人有低的回报率。对借款人而言，如果他们都拥有能带来相同期望总收入的项目，则收取更高的利率阻止的是那些想拥有更安全项目的借款人，也就是银行更希望拥有的客户。银行的利润会随着利率的上升而下降。第二条

---

* Joseph E. Stiglitz, Andrew Weiss, *The American Economic Review*, Vol. 73, No. 5 (Dec., 1983), pp. 912–927. 两位作者分别属于：普林斯顿大学经济系，哥伦比亚大学和贝尔实验室。本文仅代表作者个人观点，而非代表贝尔实验室和贝尔系统的观点。非常感谢美国国家科学基金会的资助。我们非常感谢 Henry Landau 在数学推导方面的宝贵建议。

理由是更高的利率会促使借款人从安全的项目转向更具风险的项目（激励效应）。这种转变的发生是因为对那些更有可能失败的项目，为其支付利息的可能性也更小。因此增加利率的成本越低，则违约的可能性越高。

如果用一个足够一般性的模型来解释如下所有情况——一旦坏结果发生委托人就采取终止委托—代理关系的措施，这样的模型无法体现每种市场所独有的特点，因此我们把注意力集中在分析为什么银行拒绝向违约者贷款。在本文的第六部分，我们将用相同的分析方法来解释为什么工人会被解雇。在这两个例子中，被终止的代理人都要好于（将在下文给予定义）委托人雇佣的新的代理人。

即使对合同的条款施加如下的限制，这些结论仍然成立。首先，实施合同的每一条款必须是对一方有利的：如果合同要求不再向违约者提供信贷，则银行必须发现取消信贷是有利的；如果归还贷款的借款人信贷增加，要么是银行要么是借款人必须发现增加贷款是有利可图的。其次，合同的激励效应必须考虑到其他代理者的反应。例如，如果借款人在被自己的贷款人拒绝提供贷款后可以从其他的银行借到贷款，那么“拒绝信贷”的激励效应就应该是借款人由于需要向其他贷款人借贷而导致的福利损失，而不是由于其被驱逐出信贷市场而导致的福利损失。

一、模型

我们首先假设借款人和贷款人都是风险中立的，所有的借款人都是同质的，[1] 并且借款人只存在两期（在我们的分析过程中会清楚地发现我们的结论可以被扩展到借款人可以存在很多期的情况。在文章的第四部分，我们将分析借款人是风险厌恶型的情况）。在每一期，借款人选择投资一个项目。两期可供选择的项目可能不同。我们进一步假设：如果借款人有能力偿还贷款，他必须偿还。[2] 银行向新的客户提供的贷款合同包括以下要素：[3,4]

---

1 该假设使我们能集中关注激励问题，在我们 1981 年的文章中，我们分析了信贷市场上的逆向选择问题。

2 其他一些关于信贷市场的模型，如丹尼尔·本杰明（Daniel Benjamin，1978）的，都假设当贷款的数额高于抵押品的价值（对借款人的价值），借款人就会“携款潜逃”。在我们的模型中，我们假设有一个强制执行的体系来杜绝这样的行为。然而，因为我们的主要结论是当存在对贷款的过度需求，借款人是可以在投资上获得正的利润的情况下成立的，所以，当贷款的价值超过抵押品价值就违约的策略总体而言并不是借款人利润最大化的策略。参见 Franklin Allen（1981）。

3 有三类新的客户：第一期借款人、前一期从其他银行借钱的第二期借款人、前一期被取消信用额度的第二期借款人。我们将先分析提供给新客户的借款合同。在第三部分，我们将证明均衡时，银行将不会向前一期没有向它们借款的第二期借款人提供贷款。

4 我们假设项目的“成功”和“失败”是可以观察到的，但观察不到回报的多少。参见下面的讨论。

r_1 = 第一期贷款的利率；

r_2 = 第一期贷款偿还了的借款人在第二期所需要支付的利率；

r_2^f = 第一期贷款违约了的借款人在第二期所需要支付的利率；

γ^s = 第一期贷款偿还了的借款人在第二期从银行那里得到贷款的概率；

γ^f = 第一期贷款违约了的借款人在第二期从银行那里得到贷款的概率。

向量 $C \equiv \{r_1, r_2, r_2^f, \gamma^s, \gamma^f\}$ 完整地描述了一个贷款合同。

可以很容易证明，在很多情况下，银行会向偿还了前期贷款的借款人再提供贷款，因此为了简化符号，我们假设那些情况的成立，并使 $\gamma^s = 1$。

还有一组很重要的控制条件，在将其省略前，我们需要做一些说明：通过改变对抵押品的要求，银行可以影响被实施项目的风险的高低（类似地，它可以要求更高的资产—负债比）。

由于我们考虑的是对贷款人损失很大的违约情况，所以，我们假设项目的大小和借款人拥有的财产使得贷款的金额总是超过了抵押品的金额。很明显，如果抵押品的价值等于或者大于贷款的金额，则贷款是无风险的，其利率应该和美国国债利率一样。假设存在一定的违约概率并且借款人违约时银行会受到损失应该是合理的。为了简化分析，我们忽略银行对抵押品的要求。[5]

银行无法直接观察到借款人选择的项目，[6] 但知道合同的条款是如何影响借款人的行为的（银行知道每个潜在借款人的年龄和经历）。

在每一期，潜在的借款人向银行申请贷款，该银行提供的合同使得借款人的期望收益最大化，而潜在借款人则选择能让该合同回报最大化的投资项目。银行明白这点，它们选择能够让自己期望利润最大化的合同条款（这是一个标准的间接控制问题）。我们的目的是找出均衡贷款合同具有的特点。尤其是我们证明了或许存在一个纳什均衡，在该均衡下银行拒绝向不偿

5 可以用几种方法来证明这个简化的合理性。例如，如果借款人在每一期都把所有财富消费掉，除了通过贷款获得的资产，没有抵押品可以提供。此外，在很多情况下，对抵押品的要求的增加会导致银行回报率的降低。例如，我们 1981 年的文章证明了：如果借款人持有不同的财富，但有相同的效用函数，且其效用函数都是递减的绝对风险规避型的，那么抵押品要求的提高将使对风险规避的借款人退出，从而会降低期望的回报率。Hildegard Wette（1983）已经证明了即使借款人是风险中立的，抵押品要求的提高也会导致逆向选择问题。我们的公式通过引入一个固定的抵押品水平所做的修正是非常直观的。

6 如果能观察到项目选择，则贷款合同可以简单地规定实施何种项目，也就不存在激励问题了。当然，银行常常对借款人选择某个项目有一定的控制权——我们只关注其余借款人的选择。如果银行有完全的控制权，则银行和借款人之间就没有区别了。根据现行的银行法，万一出现违约的情况，银行会没收它有权没收的那部分借款人的财产。

还其第一期贷款的借款人提供贷款。

为简化分析，我们关注项目只有两个可能结果的情况：回报为 R 时的成功和回报为零时的失败。回报为 R 的项目在第 i 期被实施并取得成功的概率为 $p_i(R)$ 。明显的有 $p'_i(R) < 0$ ，因为人们不会选择回报和成功率都相对较低的项目。[7] 同时，我们假设 $p_1(R) \leqslant p_2(R)$ ：年长的人可以从他们过去的经历中吸取经验教训。[8]

首先我们假设没有银行会借钱给竞争对手的违约客户，并证明在一定条件下所有在第一期违约的借款人在第二期都借不到钱。然后我们证明在均衡时，合同的设定将使被自己的银行拒绝贷款的违约者不能从其他地方获得贷款。在这小节，我们分析的重点是均衡时存在对贷款的超额供给的情况。在下一部分中，我们考虑银行为争取借款人和存款人而竞争的情况。

借款人：借款人（给定关于银行策略的假设，尤其是银行不向其他银行的违约客户提供贷款的假设）在第一期选择 R 的一个值（R_1）；如果他们在第一期成功了，则在第二期选择 R 的一个值（R_2）；如果他们在第一期不成功但设法取得了贷款，则在第二期选择 R 的一个值（R_2^f）。借款人选择这些投资使得期望利润（的贴现值）π 最大化：

$$\max_{\{R_1,R_2,R_2^f\}} \pi = p_1(R_1)[R_1 - (1 + r_1)] + \delta p_1(R_1)p_2(R_2)[R_2 - (1 + r_2)] + \delta\gamma^f[1 - p_1(R_1)]p_2(R_2^f)[R_2^f - (1 + r_2^f)] \quad (1)$$

此处 δ 为未来回报的贴现率（假设对借款人和贷款人都一样）。我们现在来分析项目选择是如何受借贷成本的影响的。定义：

$$\pi_2(r) \equiv \max_R p_2(R)(R - (1 + r)) \quad (2)$$

为借款人可以从第二期贷款中获得的期望利润；以及：

$$x \equiv (1 + r_1) - \delta(\pi(r_2) - \gamma^f\pi(r_2^f)) \quad (3)$$

此处的 x 表示偿还了第一期债务的借款人需要向银行支付的总利息成本。

7　如果失败时候的回报为正，我们的分析也不会有本质的变化。类似地，我们可以很容易地将分析扩展到回报是随机变量的情况。那么对风险中立的借款人而言，他们关心的是如果工程不违约时的回报均值（对应于我们模型中的 R）和违约的概率。对贷款人而言，其关心的是违约的概率和违约时的回报均值。在这种情况下，观察到某一水平的回报就无法揭示太多的信息（在正文所讨论的情况中，观察到产出的水平就能知道关于借款人所选择的项目的所有信息）。然而，这种情况的计算非常复杂，因为破产的概率一般都是银行所要求的利率的函数。

8　做这个假设是为了让我们的结论更有意思。违约者所选的项目不比新申请者的项目更具有吸引力，则其贷款申请会被拒绝，证明这点是很简单的。我们将给自己更难的任务——证明当违约者有更好的项目时该结论也成立。

$1+r_1$是偿还的利率，而$\delta\ (\pi\ (r_2)\ -\gamma^f\pi\ (r_2^f))$是偿还了贷款的借款人取得了更好的借贷条件的贴现价值。成本 x 是银行在第一期的关键控制变量，并完全决定了借款人在第一期的投资，R_1。

我们可以将式（1）重写为

$$\pi = p_1(R_1)[R_1 - x] + \delta\gamma^f\pi_2(r_2^f) \qquad (1')$$

注意到 R_1 依赖于 x 的值，而 R_2 和 R_2^f 分别只依赖于 r_2 和 r_2^f的值：

$$\partial\pi/\partial R_1 = p'_1(R_1 - x) + p_1 = 0; \qquad (4)$$

$$\partial\pi/\partial R_2 = \delta p_1[p'_2(R_2 - (1 + r_2)) + p_2] = 0;$$

$$\partial\pi/\partial R_2^f = \delta\gamma^f[1 - p_1][p'_2(R_2^f - (1 + r_2^f)) + p_2^f] = 0;$$

对式（4）做隐函数微分，

$$dR_1/dx = p'_1/[p''_1(R_1 - x) + 2p'_1] > 0; \qquad (5)$$

$$dR_2/dr_2 = p'_2/[p''_2(R_2 - (1 + r_2)) + 2p'_2] > 0;$$

$$dR_2^f/dr_2^f = p_2^{f'}/[p_2^{f''}(R_2^f - (1 + r_2^f)) + 2p_2^{f'}] > 0$$

上面几式的符号可以由二阶条件和p'_1、p'_2 以及 $p_2^{f'}$为负值得到。

贷款人：在给定上述借款人会对其条款作出反应的条件下，银行选择使其利润最大化的合同条款。用 $Y=\nu L$ 表示银行的期望利润，L 为银行借出的贷款数，ν 为其每笔贷款的期望利润。

$$\nu = p_1(R_1)(1 + r_1) - \rho^* + \delta p_1(R_1)[p_2(R_2)(1 + r_2) - \rho^*]$$
$$+ \delta\gamma^f(1 - p_1(R_1))[p_2(R_2^f)(1 + r_2^f) - \rho^*] \qquad (6)$$

此处ρ^*为银行每美元贷款的成本。争取存款人的竞争迫使ρ^*达到使 $\nu=0$ 的值。当ρ^*为该值时，使银行利润最大化的合同也让每个借款人得到的利润最大化。[9] 从而ρ^*为无风险利率，我们的模型使我们能将ρ^*处理为内生变量。

将$\nu = 0$代入$\partial Y/\partial z = L\partial\nu/\partial z + \nu\partial L/\partial z$即可得到该结论，此处$z$代表银行的任何选择变量。直观地，如果所有的银行都使从每笔贷款得到的利润最大化，而每笔贷款的利润都为零，那么没有最大化其每笔贷款的利润的银行将承受损失。

接下来，我们讨论每个银行都提供使 v 最大化的合同$\hat{C}$，而银行向存款

9 观察到银行处于另外的经济环境时，也可以得到类似的结论是很重要的。例如，如果银行在一些客户数量固定的地区是垄断者，但其可以以ρ^*的成本从国际市场上获得资金，那么银行的利润最大化使得 v 最大化，而在这样的市场中 v 可能为正值。

人支付的利率会导致对可贷资金的超额需求。换而言之，借款人获得$\hat{C}$，存款人获得利率$\rho(\hat{C}|\nu=0)$，此处$\rho(\hat{C}|\nu=0)$代表当银行提供的合同为$\hat{C}$，并且对存款人的竞争导致对其支付的利率上升到使$\nu=0$时，存款人获得的利率。$L^D(\hat{C})>L^S(\rho(\hat{C}|\nu=0))$，此处$L^S$和$L^D$分别代表贷款的供给和需求[10]（为了简化符号，在不会引起误解的时候，我们将用$\rho(C)$表示$\rho(C|\nu=0)$）。我们再来描述提供给风险中性的借款人的合同的特点。由前文知R_2只由r_2决定，定义：

$$S_2(r_2)\equiv p_2R_2-\rho^* \tag{7a}$$

为从第二期贷款中可获得的期望社会回报，以及

$$\nu_2(r_2)=p_2(1+r_2)-\rho^* \tag{7b}$$

为银行从第二期贷款中可获得的期望回报，用$\{r_1^*, r_2^*, r_2^{f*}, \gamma^{f*}\}$表示能最大化$\nu$的值，而用$\nu^*$表示$\nu$的该最大值。

用$\hat{r}$表示最大化了$\nu_2(r)$的r值，用$\tilde{r}$表示最大化了$S_2(r)$的r值。由$S_2(r)$和$\pi_2(r)$的定义即可得$1+\tilde{r}_2=0$。当$r_2=-1$，$\partial S_2(r)/\partial R_2\equiv\partial\pi_2(r)/\partial R_2$使$\pi_2(r)$最大化的借款人也可以满足$S_2(r)$最大化。因此在第二期，被要求任一正利率的借款人都会选择比社会最优[11]更有风险的项目。因为只有在项目成功了才会偿付贷款，投资决策被扭曲为偏好更具风险的项目，这类项目更不易成功，但一旦成功回报更高。

我们可以用式（7a）、式（7b）、式（2）和式（3）将式（6）重写为：

$$\nu=p_1x-\delta\gamma^f p_1\pi_2(r_2^f)+\delta p_1S_2(r_2)+\delta\gamma^f(1-p_1)\nu_2(r_2^f)-\rho^* \tag{8}$$

式（8）将银行每笔贷款的利润表示为x，r_2，r_2^f和γ^f的函数。r_2和r_2^f必须满足可实施约束条件，即降低利率也必然对银行和借款人都是不利的，从而$r_2^*\leqslant\hat{r}_2$和$r_2^{f*}\leqslant\hat{r}_2$。

由于x是r_1的线性函数，而r_1是无界的，银行对r_2，r_2^f和γ^f的选择不限制其对x的选择。R_1只是x的函数，我们可以将银行最优问题的一阶条件写作：[12]

10　读者可能已经观察到我们还没有定义在这个市场上什么是均衡，甚至还怀疑稳态均衡的存在；我们也没有讨论如果$L^D(\hat{C})<L^S(\rho(\hat{C}|\nu=0))$时银行提供的合同。为了使对银行的最优化问题的分析尽可能简单，我们将这些问题的讨论放在文章的第三部分。

11　此处这个概念是指能够最大化期望净国民产值的项目，这种用法是不太严格的。

12　在推导（9a）时，我们用到了$\pi_2+\nu_2=S_2$。

$$p_1 + p'_1(dR_1/dx)[x + \delta S_2(r_2) - \delta\gamma^f\pi_2(r_2^f) - \delta\gamma^f\nu_2(r_2^f)] = 0; \quad (9a)$$

$$\partial\nu/\partial r_2^f = \delta\gamma^f[-p_1\pi'_2(r_2^f) + (1-p_1)\nu'_2(r_2^f)] \geqslant 0; \quad (9b)$$

$$\partial\nu/\partial r_2 = \delta p_1 S'_2(r_2) \geqslant 0 \quad (9c)$$

式（9b）和式（9c）中的不等式只有当可实施约束条件对相应的利率成立时才成立。我们现在来分析当借款人为风险中性时，式（9a）至式（9c）推导出的均衡合同的结构是什么样的。

1. 第二期利率

从式（9c）可推出 $r_2 = \tilde{r}_2$，所以对成功的借款人所要求的第二期利率是社会最优的，即 $1 + r_2^* = 0$，可实施约束条件此时不无约束力（正如我们在第四部分证明的，如果借款人是风险规避的，则 $1 + r_2^* > 0$）。

另一方面，由式（9b）可以得到，如果缺少可实施约束，在第一期违约的借款人将被要求支付能最大化 $\delta\gamma^f[(1-p_1)\nu_2(r_2^f) - p_1\pi_2(r_2^f)]$ 的利率。但是，由于 $\pi'_2(r_2^f) < 0$，可实施约束总是成立。因此违约者被要求支付的利率为 $\hat{r}_2$。[13]可实施约束确保了他们不被收取更高的利率。在这一点看来，利率的增加会增加对第二期贷款的回报，并使公司承受更高的第一期利率，同时确保公司不会改变第一期实施的项目的风险。这些结论有一个更深的含义：获得贷款的违约者选择的项目比成功的借款人更具风险。他们的失败率更高，不是因为他们的能力较差（假设了每个个体都是同质的），而是因为他们被惩罚而被收取了更高的利率。正是更高的利率诱使他们投资于更有风险的项目。

2. 第一期利率

从式（3）、式（7a）、式（7b）、式（9a）和 $r_2^f = \hat{r}_2$，我们得到

$$p'_1(1 + r_1)(dR_1/dx) + p_1 + \delta(dR_1/dx)p'_1[\nu_2(r_2) - \gamma^f\nu_2(\hat{r}_2)] = 0$$

正如我们已经证明了的，$\nu_2(r_2) < \nu_2(r_2^f)$，且 $\nu_2(r_2) < 0$（倘若 $\rho^* > 0$，我们将假设其成立）。与 $\hat{r}_2$ 相似，我们定义 $\hat{r}_1$ 为下式的解：

$$p'_1(1 + r_1)(dR_1/dr_1) + p_1 = 0$$

并称 $1 + \hat{r}_1$ 为（短视的）第一期银行最优利率。因为括号里的项对 $\gamma^f = 0$ 和 $\gamma^f = 1$ 都为负，其对所有的 γ^f 都为负，且 $1 + r_1^* \geqslant 1 + \hat{r}_1$。银行有能力在制定第二期贷款的条款时视第一期贷款的结果而定，这使得银行增加在第一期收

13 如果我们没有加入可实施约束，对违约者收取的利率应该等于 arg max $\{(1-p_1)\nu_2(r_2) - p_1\pi_2(r_2)\}$。用该表达式替代每个 r_2^{f*}，可以证明我们定性上的结果是不变的。

取的利率。

在我们1981年的文章中的一期模型中，银行可用于影响借款人行为的唯一工具是当期收取的利率。这些激励效应使得银行收取的利率要比当它可以控制借款人行为时收取的利率低。带意外条款的合同给了银行更多的工具来控制借款人，因此使得银行增加了其在第一期的利率。

3. 信贷限制——γ^f的值

注意到ν是γ^f的线性函数，而x的取值范围不受γ^f的限制。[14]因此：

$$\gamma^{f*} = \begin{cases} 0 \\ 1 \end{cases} \text{as} p_1 \pi_2(r_2^f) - (1 - p_1)\nu_2(r_2^f) \gtrless 0 \tag{10}$$

此处，p_1、$\pi_2(r_2^f)$、$\nu_2(r_2^f)$的取值为代入一定值的r_1、r_2、r_2^f算出，而r_1、r_2、r_2^f此时的取值为在$\gamma^f = 1$的约束条件下最大化ν的值。[15]

因此如果加入了可实施约束，当且仅当下式满足时，才有$\gamma^{f*} = 0$：[16]

$$p_1 p_2^f(R_2^f - (1 + \hat{r}_2)) > (1 - p_1)[p_2^f(1 + \hat{r}_2) - \rho *] \tag{11}$$

通过分析当$\gamma^f = 1$的例子我们可以看到对式（10）直观的解释。而从式（3）我们能发现拒绝向违约者贷款（使$\gamma^f = 0$）使银行在保持R_1不变的情况下将$1 + r_1$增加$\delta\pi(r_2^f)$。由于只有当项目成功时，利率才会被支付，更高的利率带给银行的价值为$p_1\delta\pi_2(r_2^f)$。这部分是贷款给违约者的银行成本——因为这些贷款会对银行在第一期能收取的利率带来影响，并且会引起借款人期望的行为。另一方面，$(1 - p_1)\delta\nu_2(r_2^f)$为这些贷款的期望现值。只有当向违约者提供贷款所能带来的利润多于负面的激励效应时，银行才会向违约者贷款。

因为我们只关心证明最大化ν的合同可以用$\gamma^{f*} = 0$来描述，我们来看看当$p_2(R) \equiv p_1(R)$的情况。将银行的问题视为最大化支付给存款人的利率ρ^*，而不是最大化其每笔贷款的回报ν，会使对问题的分析更简单一些。对存款人的竞争使得这些最大化问题的解都是一样的。

14　ν是γ^f的线性函数以使得γ^{f*}的值为0或1——这个结论依赖于借款人是风险中性的假设。在第四部分，我们将证明如果借款人为风险厌恶的，γ^{f*}可能有内点解。

15　因为γ^f的值为0或1，我们只需要计算$d\nu/d\gamma^f$在$\gamma^f = 1$时的值（当$\gamma^f = 0$时，r_2^f可以取任意值，因此不用计算式（10）在$\gamma^f = 0$时的取值）。

16　可以证明式（10）只有在意外的情况下才为等式。

让我们考虑次优合同 $\tilde{C}=\{r_1=r_2^f=\hat{r}_2,\ 1+r_2=0,\ \gamma^f=1\}$；[17]那么有：

$$\rho^* > \rho(\tilde{C}) = [p_2^f(1+\hat{r}_2)+\delta p_2^f(1+\hat{r}_2)]/(1+\delta) \quad (12)$$

因此，

$$p_2^f(1+\hat{r}_2)-\rho^* < 0 \quad (13)$$

因为式（11）的左边为正，如果 $p_2(R)=p_1(R)$，则银行通过拒绝向违约者贷款而使它对每个借款人的贷款回报最大化。直观的解释是，如果第二期的项目不比第一期的更好，银行可以通过拒绝向违约者贷款来低成本地提高 p_1。因为这严格增加了银行的利润，限制信贷的政策也可能是利润最大化的，即使 $p_2(R)$ 比 $p_1(R_1)$ 更大，且 $\hat{r}_2>\hat{r}_1$。

4. 总结

使银行每一美元贷款回报最大化的合同有两种形式：

（a）当借款人在第一期没有违约时，银行才会提供第二期的贷款（$\gamma^f=0$）；

（b）不管第一期借款人的表现如何，银行都在第二期提供贷款。对违约者收取比偿还了第一期贷款的借款人要高的利率。

在这两种情况下，银行在第一期收取的利率都比当银行不能将借款人在第一期的表现与第二期贷款的获得同条款联系起来时所能收取的利率要高（例如，$r_1^*>\hat{r}_1$）。如果银行通过减少贷款而获得一期项目收益，能够提高第一期的利率，并且由此带来的第一期利润的增加大于其第二期向违约者提供贷款可能获利的贴现值，那么银行将选择拒绝贷款给违约者。

二、银行间的竞争

在前一部分中，我们假设当银行拒绝贷款给一家公司，则这家公司不能找到其他的资金来源。在这一部分中，我们将分析这是否是一个正确的假设。

银行通过一个简单的机制确保违约者不能从其他地方获得贷款。通过在第一期的贷款合同中规定第一期的贷款比后面的贷款有优先权，银行可以降低第二期借款人的期望回报。尤其是，即使银行借贷给无限制的且有贷款经历的（第二期）借款人是有利可图的，如果这些借款人已有一些必须优先

17 如我们已经证明的，最优合同由 $r_1^*>\hat{r}_1$ 描述。

于新贷款获得偿还的贷款，银行也可能选择拒绝贷款。[18]

我们现在允许借款人宣告破产以消除以前的债务。此时仍可能存在一个均衡，在该均衡下，第二期贷款申请被拒绝的借款人会被其他所有的银行所拒绝。第二期贷款的最大回报为 $\nu_2(\hat{r}_2)$。在 $\gamma^f=0$ 的情况下，第一期贷款给借款人的银行将能给存款人的回报为：

$$\rho^* = \frac{p_1[1+r_1+\delta p_2(1+r_2)]}{1+\delta p_1} \tag{14}$$

如果 $\rho^*>p_2(1+\hat{r}_2)$，那么向违约者提供贷款的银行都不能吸引到存款人。[19] 为了证明该不等式在不违反使得 $\gamma^f=0$ 为银行最优的条件的情况下也成立，我们假设 $r_2=\hat{r}_2$（该限制条件只使 ρ^* 相对最优的 r_2 减少）。这已足以证明 $\gamma^{f*}=0$ 和 $\nu^*>\nu(\hat{r}_2)$ 条件成立的存在。很明显，如果 $\forall R$，$p_1(R)=p_2(R)$ 成立，则该不等式成立，因为此时 $p_1(1+r_1^*)>p_2(1+\hat{r}_2)$。

假设一个“短视”的银行会发现第二期的贷款的回报比第一期的贷款更高。即使那样，对银行而言，带意外条款的合同所带来的回报也比第二期贷款的更高。第一期贷款的收益能力取决于能否让第二期借款人的合同条款与第一期贷款偿还与否联系起来。在那些情况下，对存款人的竞争使银行拒绝贷款给违约者。[20]

三、信贷市场的均衡

迄今，我们已经证明了使银行稳态利润最大化的合同可被描述为拒绝向违约者贷款的合同。然而我们回避了讨论该合同是否是市场的均衡结果，或当银行之间需要为争取借款人和存款人而竞争时，该结果是否还成立。假设我们现在考虑 $L^D(\hat{C})<L^S(\rho(\hat{C}|\nu=0))$ 的情况，如前所述，$\hat{C}$ 为最大化 ν 的合同，$\rho(\hat{C}|\nu=0)$ 为当提供合同 $\hat{C}$，借款人和贷款人的利润都为 0 时，支付

---

18 该结论并依赖于借款人只存在两期的假设。即使代理者无限期存在，由于要求借款人按照债务发生的先后顺序偿还，则次级债务会使得对任一银行或潜在的进入者而言，贷款给竞争对手的违约客户是无利可图的。

19 $\rho^*>p_2(1+\hat{r}_2)$ 这个条件也是“拒绝向违约者贷款”成为合同中具有实施性的条款的充要条件。如果不等式成立，银行偏好于贷款给之前没有贷款经历的借款人，曾经借贷并偿还了贷款的借款人会确保其贷款合同中的续借条款的实施，而银行会确保合同中要求拒绝贷款给违约者的条款的实施。

20 读者应该记住这个讨论是假设对能使银行利润最大化的贷款合同，存在超额的贷款需求。

给存款人的回报。在这部分，我们将证明即使银行要竞争借款人，市场均衡也可能是违约者被拒绝贷款。

为了简化证明过程，我们只分析对称均衡。银行的均衡策略用S^*表示，其由以下几部分构成：贷款合同C^*、支付给存款人的利率ρ^*以及存款期（事前已经排除了存款在一期结束后就被提取的可能性）。我们将策略S^*定义为对称均衡，如果所有的银行都采取该策略，且不存在另一策略$S \neq S^*$更吸引存款人（下面的不等式（15））和借款人（下面的不等式（16））。[21]即不存在一个策略$\tilde{S}$满足：

$$\rho(\tilde{S}) \geq \rho(S^*) \tag{15}$$

和下两式中的一个

$$\pi(\tilde{C}) \geq \pi(C^*) \tag{16a}$$

对获得贷款的借款人而言，或者

$$\pi(\tilde{C}) \geq 0 \tag{16b}$$

对没有获得贷款的借款人而言，

此处$\pi(C)$是被提供合同C的借贷人的期望利润，[22]并且式（15）、式（16a）或者式（16b）为严格的不等式。

由此可以推得均衡的合同带给银行的是零利润，任何的非零利润均衡都能被将自己的一部分利润分给存款人（或借款人）的银行所打破。

我们只考虑可实施和可信的合同。我们在上文（第一部分）已经讨论过可实施性的概念。可信的策略的特征是：如果带意外条款的合同要求对第一期借款人提供一定数额的第二期贷款，则可信的策略必须能够保证获得所需要的资金。类似地，预期将把一定百分比的新贷款贷给新借款人，因而两期都向存款人提供$\tilde{\rho}$的利率的策略，只有当其总能获得那些新借款人时才是可信的。即该策略不仅在给定其他银行的策略时能吸引新的借款人，而且当其他的银行采取任何可获利的（但不一定可信的）的策略时也能吸引新的

21　该均衡概念不允许一个能吸引所有存款人，使得其他银行没有资金可以用以借贷，从而获得了所有的借款人的合同的存在来破坏该均衡。该均衡概念类似于一个没有一方支付的博弈。即不存在借款人和贷款人的联盟，当报酬完全由合同的条款决定时，该联盟是可以形成并使每个成员的福利增加的。

22　该表达式假设存款人不是被分配的。证明该假设总是成立是很容易的。直观地，分配存款人是没有意义的，因为被分配的存款人可以压低回报ρ直至或者没有超额供给或者$\rho=0$的程度，而存款人对将资金存在银行和放在家中是无差异的。

借款人。这些可信性条件只有在以下情况才能被满足：所有签订了两期带意外条款的合同的借款人和两期都将其资金的 p_2 部分进行投资的存款人相匹配。[23,24]通过加入这些可信性要求，我们可以防止稳态均衡被扩张性的银行所破坏，该银行在扩张的过程中会增加对新借款人贷款的比重，并用更高的利润来增加支付给存款人的利率。

均衡的结构依赖于 $L^D(\hat{C}) \geqslant L^S(\rho(\hat{C}|\nu=0))$ 这个条件。如果对可借贷资金的需求和供给使得当所有的银行都提供最大化 ρ^* 的合同 $\hat{C}$ 时，对贷款存在超额需求，那么，如我们已经证明的，最大化 ν 等价于最大化 ρ^*。第一部分和第二部分的结论在这里也适用：在合理的条件下，市场均衡为 $\gamma^f=0$。然而，由于

$$L^D(\hat{C}) > L^S(\rho(\hat{C}|\nu=0))$$

不仅所有的违约者都被拒绝提供贷款，而且一些第一期借款人也被拒绝了。[25]

第二个比较有意思的是，当银行竞争借款人时：$L^S(\rho(\hat{C}|\nu=0)) > L^D(\hat{C})$。这个均衡也可以用不向违约者提供贷款来描述。用 π^0 表示借款人的均衡利润水平，银行竞争借款人，所以均衡合同是下式的解：

$$\max\rho(C)\text{，约束条件为 } \pi(C) \geqslant \pi^0\text{。}$$

我们构造拉格朗日函数 $H=\rho(C)-\lambda(\pi^0-\pi(C))$，在均衡时：

$$\partial H/\partial r_1 = \partial\rho(C)/\partial r_1 + \lambda d\pi/dr_1 = 0$$

$\lim_{C\to\hat{C}}\partial\rho(C)/\partial r_1=0$，和 $\lim_{C\to\hat{C}}d\pi/dr_1<0$。假设在接近 $\hat{C}$ 的合同处市场出清。那么，因为在均衡合同时，$\partial H/\partial r_1=0$，$\lambda$ 将很小。均衡合同 C^* 用 $\gamma^f=0$ 来表示，如果下式成立：

23 需要注意的是没有和长期存款人联系起来的长期借款人将无法获得第二期贷款，除非银行能够在第二期吸引到新的存款人（银行的贷款数增加了）。这些存款人是不可得的，如果其他一些银行在后一期开始的时候选择类似的扩展政策。另一方面，如果发现“承诺的”利率是来自于第二期向新的无经验借款人贷款的银行，存款人也不会投资两期。没有理由让存款人相信会有新的借款人。结果，可信的策略的概念排除了通过构造庞氏博弈（Ponzi-type Game）来增加借款人或存款人回报的策略，在该类型的博弈中，银行通过向新借款人贷款获取利润，并连续不断地用新贷款来弥补它对有贷款经历借款人受到的损失。可信性要求防止了银行被迫依赖新贷款（或新存款人）来为对借款人或存款人的许诺增彩的情况。

24 需要注意的是我们不需要对存款人的寿命做任何的限制，因为即使存款人只存活一期，他也可以向银行做多期的承诺，这些承诺迫使他的后代在出售资产之前或者将这些资产在二级市场上出售等待。

25 对于第一期配给的更全面的讨论，请参见我们1981年的文章。一般地存在一个市场出清的利率使得对贷款的需求等于供给。银行可以通过降低它所收取的一个或更多的利率来增加它的回报。

$$\left.\frac{\partial H}{\partial \gamma^f}\right|_{\gamma^f=0} = \frac{\partial \rho(C)}{\partial \gamma^f} + \frac{\lambda d\pi}{d\gamma^f} \leqslant 0 \tag{17}$$

考虑一组使$\partial \rho/\partial \gamma^f < 0$在$C = \hat{C}$成立的参数。那么对$\lambda$足够小的值，

$\lim_{C\to\hat{C}}\partial H/\partial \gamma^f|_{\gamma^f=0} < 0$。因此，如果市场在接近$\hat{C}$合同处出清，则$\lambda$的值小且均衡合同为$\gamma^f = 0$。

假设在$\hat{C}$处，存在对资金的超额供给。银行通过改变合同，增加借款人的期望回报π（C）来竞争借款人。它们试图通过最小化支付给存款人的成本来达到此目的。银行改变其合同的方法使得π（C）的增加在很大程度上对应于ρ（C）的下降。然而，在$C = \hat{C}$，r_1的降低对ρ（C）只存在二阶影响（r_1有内点解）。所以通过降低r_1可以低成本地增加π（C）。另一方面，由于γ^f为边界点，增加π（C）所需要的γ^f的增加会对ρ（C）的减少有一阶影响。

如果L^S（ρ（$\hat{C}$））$> L^D(\hat{C})$，即如果在合同$\hat{C}$处存在对可贷资金的超额供给，均衡合同则为γ^f的值为［0，1］中的任意取值：即使所有的第一期借款人都获得贷款，一些违约者也会被拒绝贷款。对第一期借款人的配给是应急性的而不是任意的。[26]

从最大化H的一阶条件，我们可知对$\gamma^{f*} > 0$，均衡为$\partial H/\partial \gamma^f = 0$，$\lambda > 0$，且

$$\frac{d\pi/dr_1}{\partial \rho(C)/\partial r_1} = \frac{d\pi/d\gamma^f}{\partial \rho(C)/\partial \gamma^f} \tag{18}$$

式（18）说明了如果市场均衡是一个内点解，对每个工具而言，边际收益（借款人增加的利润）的每一美元边际成本是一样的。

---

26 到目前为止，对L^S（ρ（$\hat{C}$））$< L^D$（$\hat{C}$）的情况，我们只分析了能最大化ρ^*的合同。然而，在一些有点“反常”的情况下，即使L^S（ρ（$\hat{C}$））$i < L^D$（$\hat{C}$），仍可能存在市场出清的均衡。假设存在一些合同$\tilde{C} \neq \hat{C}$，使得L^S（ρ（$\tilde{C}$））$= L^D$（$\tilde{C}$），并对任何$C \neq \tilde{C}$，有当ρ（C）$> \rho$（$\tilde{C}$），则π（C）$< \pi$（$\tilde{C}$），如果π（C）$> \pi$（$\tilde{C}$），有ρ（C）$< \rho$（$\tilde{C}$）。那么$\tilde{C}$即为一个均衡，因为任一其他的合同都会使借款人或者存款人受损。此时即使在合同$\hat{C}$下，$\gamma^f = 0$，在合同$\tilde{C}$下，$\gamma^f = 1$，违约者得到贷款。在那种情况下，同时存在一个市场出清的均衡和非市场出清的均衡。当支付给存款人的利率降低导致资金供给的增加，或者增加借款人利润的政策使得对贷款的需求减少时会出现上述的情况。这些有些反常的价格效应必须足够大，以抵消增加的γ^f会增加贷款需求而导致的直接影响。当L^S（ρ（$\hat{C}$））$< L^D$（$\hat{C}$）时，多个均衡的存在对我们的均衡的定义是敏感的。如果我们要求只有竞争合同才吸引存款人，则L^S（ρ（$\hat{C}$））$< L^D$（ρ（$\hat{C}$）时的唯一对称均衡为合同$\hat{C}$。

图 1 展示了一些均衡结果。我们在 r_1，γ^f的空间中保持 $1+r_2=0$ 和 $r_2^f=\hat{r}_2$,画出了借款人和银行的等利润线。向上倾斜的为借款人的等利润线，向下和向右移动都增加借款人的利润。如我们所证明的，ν 在 $\gamma^f=0$，$r_1=r_1^*$ 时为最大值。如果对借款人的均衡回报对应于和 r_1 轴相交于 r_1^* 的等利润线，例如 π_1，则在均衡时，$r_1=r_1^*$，$\gamma^f=0$。该均衡存在对贷款的超额需求，然而，降低借款人利润的合同也会降低贷方的回报 ν。

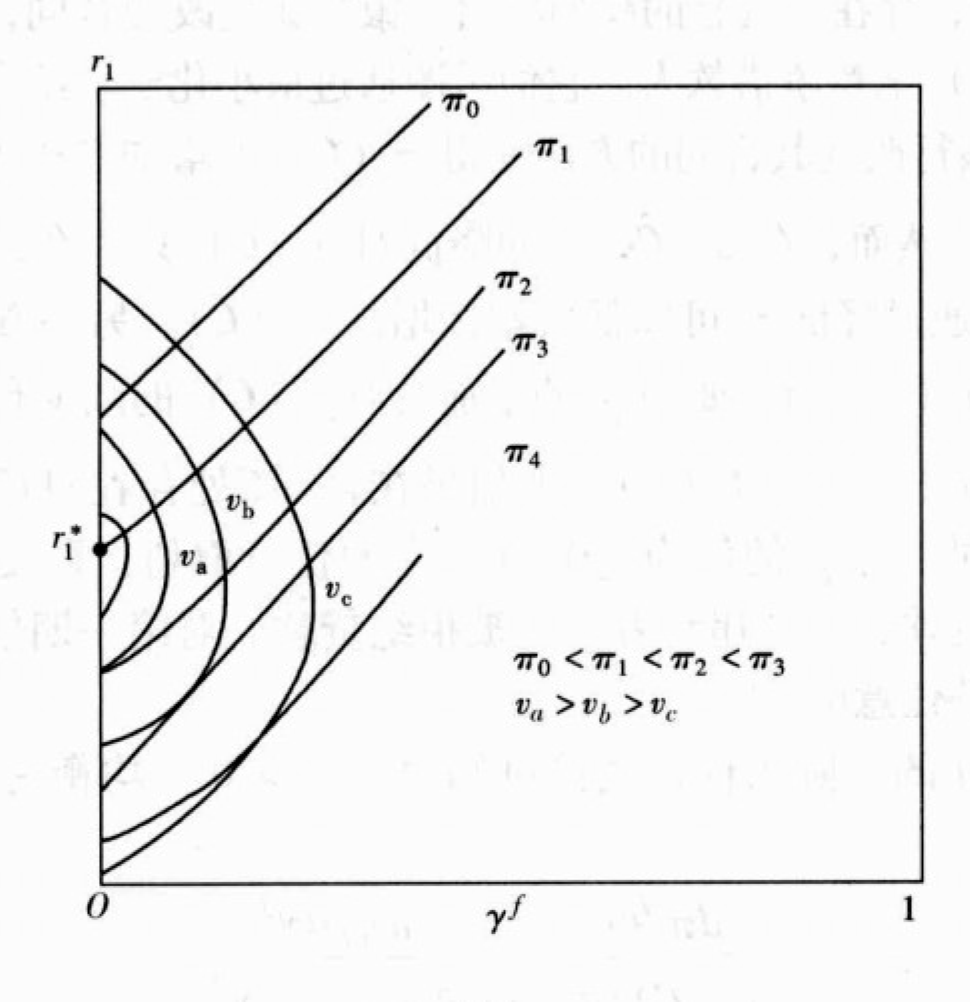


图 1

如果借款人的均衡回报在大于 π_1 的等利润线上，但不同银行的等收益曲线相切，如 π_2，那么均衡的合同为 $\gamma^f=0$，$r_1<r_1^*$。银行通过降低利率来竞争借款人。

当均衡等利润曲线，如 π_3 或 π_4，和对应的银行的等收益曲线 ν_c 有内点相切点时，均衡合同为式（18）所决定。

四、风险厌恶的借款人

目前为止的分析都假设借款人和贷款人都是风险中性的。就银行对单个贷款——这是我们所关心的——的偏好而言，该假设是比较合理的近似。然而，对借款人一生收入贴现值的最大化假设可能在很大程度上扭曲了借款人的偏好。如果我们假设借款人具有更一般性的、可加可分性的冯·诺依曼—摩根斯坦效用函数（von Neumann-Morgenstern utility function），上面得到的

一些结论将不再成立了。

均衡仍由以下方法得到：在借款人选择 R_1、R_2 和 R_2^f 以最大化其期望效用$\hat{U}$的限制条件下，银行选择一个带意外条款的合同来使 $\rho(C)$ 最大化。

$$\hat{U} = p_1[U(R_1 - (1 + r_1)) + \delta p_2 U(R_2 - (1 + r_2)) - \delta\gamma^f p_2^f U(R_2^f - (1 + r_2^f))] + \delta\gamma^f p_2^f U(R_2^f - (1 + r_2^f)) \tag{19}$$

此处 U 为借款人在每一期的效用函数。U_1 表示第一期合同带给借款人的效用，U_2^s 和 U_2^f 分别为第二期合同带给其第一期合同成功和失败了的借款人的效用。

因为借款人选择 R_1 使$\hat{U}$最大化，

$$d\hat{U}/dR_1 = p'_1[U_1 + \delta p_2 U_2^S - \delta\gamma^f p_2^f U_2^f] + p_1 U'_1 = 0 \tag{20}$$

对任意的$\{\gamma^f, r_1\}$，都有一个 R_1 满足式（20），因此：

$$\left.\frac{d\gamma^f}{dr_1}\right|_{R_1} = -\frac{p'_1 U'_1 + p_1 U''_1}{\delta p'_1 U_2^f p_2^f} < 0$$

$$\left.\frac{d^2\gamma^f}{dr_1^2}\right|_{R_1} = \frac{2p'_1 U''_1 + p_1 U'''}{\delta p'_1 p_2^f U_2^f}$$

如果由递减的绝对风险厌恶系数决定的 $U''' > 0$，那么保持 R_1 值不变，r_1 是 γ^f的递减的凹函数。我们现在来证明 γ^f在 0 到 1 之间的取值有可能最大化 ν。在图 2 中，我们画出了对应于选取（最优值）$R_1 = R_1^*$ 的企业的$\{\gamma^f, r_1\}$。其是向下凹的。R_1^* 的值固定不变，从式（8）可得银行在 r_1，γ^f 的空间中的等利润曲线是线性的。因此，当 γ^f的取值在 0 和 1 之间时能得到 ν 的最大值。

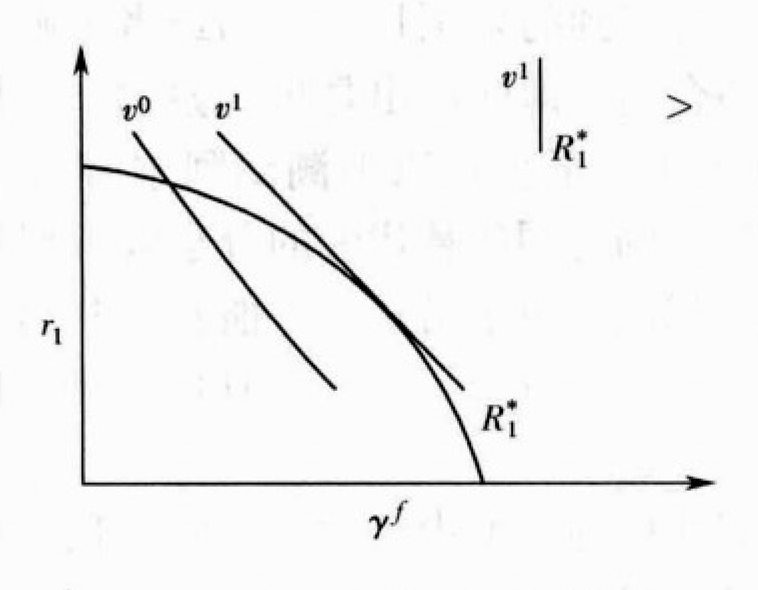


图 2

另一方面，如果借款人是风险中性的，使得 $U''_1 = U'''_1 = 0$，那么代表某个投资决策的 γ^f 和 r_1 的值的图形将是线性的。对应于该投资的银行的等利润曲线也是线性，那么除非两条线的斜率正好是相同，否则 ν 在边界点 $\gamma^f = 0$ 或者 $\gamma^f = 1$ 处取最大值。

讨论到这里已经清楚地说明了市场的均衡可能是银行将贷款随机地分配给以前违约过的借款人，即使银行同时也在相互竞争并争取新的借款人。如果最大化 $\rho(C)$ 的合同 $\hat{C}$ 为 $0 < \gamma^f < 1$，$L^S > L^D$ 时出现，银行为了争取借款人会降低 r_1，提高 γ^f。

$1 + r_2 = 0$ 的结论就像 $\gamma^f = 0$ 或者 1 的结论一样，很依赖于借款人是风险中性的假设。如果借款人是风险规避的，在 $1 + r_2 = 0$ 的领域内将 r_2 增加一小点，对借款人的效用的影响相对较小。银行可以增加 r_2，同时降低 r_1，这样可使 R_1 不变。因为借款人是风险规避的，r_2 的增加相对要大于 r_1 的减少，所以 ν 会增加，因此 $1 + r_2^* > 0$。

五、均衡的负债—权益比

这篇文章讨论的是激励问题。银行不能完全地监督借款人的行为，它们只能在每期期末时发现项目是否是成功的。因此，为了引导借款人实施安全的项目，银行得依赖于合同的两条关键条款：未来期的贷款的可获得性和收取的利率。

虽然我们已经讨论这种固定费用型的带意外条款的合同所带来的许多结果，但我们仍没有解释为什么这种借贷合同明显地会导致困难（不效率），但还是被银行所采用，以及为什么银行不用分红制的合同。在前面给出的基本模型中，由于观察到的是项目的成功或者失败而不是产出，我们讨论的合同只是可行的合同。即使产出是可观察的，使用这种借贷合同也是比较好的。大量的文献（至少可以追溯到阿尔弗雷德·马歇尔时期的）证明了作为我们所讨论的合同的替代者的分享收益型的合同（如分成制或权益融资）存在激励问题。由于借贷人不能获得其额外的努力（或者其他的投入）所带来的全部（边际）回报，这样的安排会使代理人的积极性不高。

在任何个人努力不起作用的模型中，均衡合同都是纯权益合同。另一方面，如果成功的概率不是由代理人而是回报来决定的（他无法选择方法），并且如果成功是努力的函数，以及代理人是风险厌恶的，那么均衡

合同就是纯权益借贷合同。如果公司经理既需要选择投入的努力又要选择影响破产概率的方法，则存在一个均衡的负债—权益比（debt-equity ratio），该负债—权益比平衡了两个激励问题所带来的（边际）无谓损失（deadweight losses）[27]（在一定条件下对均衡负债—权益比的推导，参见我们1980年的论文）。因此莫迪格利安尼—米勒定理（Modigliani-Miller）没有考虑道德风险显然是非常重要的。

六、解雇员工

在本章导言（一、模型以前的内容）部分，我们曾提到我们提供的分析方法可以被用来解释很多种委托—代理关系中的类似于带意外条款的合同现象。尤其是我们的模型可以被一般化，不仅能帮助我们理解为什么银行会拒绝贷款，还能帮助我们理解为什么公司会解雇员工，地主为什么要终止和承租的佃农的租约，以及为什么保险公司拒绝续签保险单。

为了证明之前提出的方法如何能够被用到其他市场，我们来考虑劳动力市场上的一个标准模型。假设工人是风险中性的，且只工作两期，在第 i 期他们投入 e_i 的努力，这会影响到他们为之工作项目的成功概率。成功项目的价值为 vQ_i，此处 Q_i 为工人的产出，v 为每单位产出的价格。我们假设 vQ_i 独立于每个工人或其努力程度，公司之间的竞争迫使 v 调整到使每个公司利润为0时的值。工人和公司都是风险中性的。同信贷市场例子一样，为了使问题更有意思，我们假设 $vQ_2 \geqslant vQ_1$。公司向工人提供的带意外条款的合同由以下要素构成：

ω_1 = 新雇佣的工人获得的工资，如果他的项目成功了；

ω_2^s = 工人收到的第二期的工资，如果其第一期和第二期的项目都成功了；

ω_2^f = 工人收到的第二期的工资，如果其第一期项目失败了而第二期项

27 有时人们建议通过要求“代理人”买债券，或者借债的时候提供抵押品来消除激励问题。在我们1981年的文章中我们证明了为什么这些方案常常不够。许多人并没有抵押品（或者没有资金购买债券），而只向那些有抵押品的人提供资金的限制又不是有利的（或社会有效的）：首先，他们可能不是资金的成功的管理者；其次，拥有大量流动资金的人可能是以前实施了高风险项目的人，而这些人在未来也倾向于选择高风险的项目；再次，如果更富的人的风险规避程度更低，那么增加抵押品的要求银行客户的平均风险偏好，因此而可能降低银行的利润。由于所有的这些原因，对抵押品的要求可能设置在一个足够低的水平，使得银行在违约出现时的损失会超过抵押品，而银行可能会对第一期借款人实行定量配给。与此同时，值得注意的是，由于第一期项目失败，第二期借款人的财富更少，因此使用抵押品要求会有效地将违约者从信贷市场上排除出去。

目成功了；

γ = 如果工人第一期项目失败了，其在第二期被雇佣的概率（同信贷市场上一样，证明之前的项目成功了的代理人将不会被终止合作是没有多大意思的）。

公司不能观察到个人的努力程度，只观察到项目是否成功。我们还假设公司不能通过罚款（支付负的工资）来处罚失败。因此，由于工人是风险中性的，如果项目失败，向工人支付零工资也是没有多大意思的。我们通过假设公司和工人的贴现率都为 0 来进一步简化表达式，公司的生产为规模报酬不变的，对个人而言单期的效用函数为 $\omega_i - e_i$。

公司选择$\{\omega_1, \omega_2^s, \omega_2^f, \gamma\}$来最大化

$$\pi = p_1(e_1)[vQ_1 - \omega_1] + p_1(e_1)p_2(e_2^s)[vQ_2 - \omega_2^s] + \gamma(1 - p_1(e_1))p_2(e_2^f)[vQ_2 - \omega_2^f]$$

此处 e_2^s 和 e_2^f 分别表示其第一期项目成功或者失败了的工人在第二期付出的努力程度。

工人选择来使其一生的期望效用最大化的 e_1、e_2^s、e_2^f，[28]而公司在工人选择的约束下使 π 最大化：

$$U \equiv p_1(e_1)\omega_1 - e_1 + p_1[p_2(e_2^s)\omega_2^s - e_2^s] + (1 - p_1)\gamma[p_2(e_2^f)\omega_2^f - e_2^f]$$

为了招到工人，必须有 $\max U \geqslant \hat{U}_1 + \hat{U}_2$，此处$\hat{U}_i$ 为一个工人在第 i 期可以从其他地方获得的工资所带来的效用。

为了在第二期留住工人，

$$p(e_2^s)\omega_2^s - e_2^s \geqslant \hat{U}_2; p(e_2^f)\omega_2^f - e_2^f \geqslant \hat{U}_2$$

为了简化表达式，我们假设效用约束方程不起约束作用。使

$$U_2^s = p_2(e_2^s)\omega_2^s - e_2^s; U_2^f = p_2(e_2^f)\omega_2^f - e_2^f$$

那么：

$$\partial U/\partial e_1 = p'_1(e_1)[\omega_1 + U_2^s - \gamma U_2^f] - 1 = 0$$

使 $y = \omega_1 + U_2^s - \gamma U_2^f$，并注意到 e_1 只是 y 的函数。将 π 重写，我们发现：

$$\pi = p_1(e_1)[-y - e_2^s + \gamma e_2^f + vQ_1 + p_2^s vQ_2] + \gamma p_2(e_2^f)[(1 - p_1(e_1))vQ_2 - \omega_2^f]$$

$$\delta\pi/\delta\gamma = p_1(e_1)e_2^f + [1 - p_1(e_1)]p_2^f(e_2^f)vQ_2 - p_2^f(e_2^f)\omega_2^f$$

由于 y 的取值范围不受 γ 的限制，而 π 是 γ 的线性函数，我们发现当效用

28　我们已经将努力程度标准化了，使得每单位收入的变化或者每单位努力程度的变化对效用的影响是一样的。我们的结论都不受标准化的影响。

约束不起作用时，γ 或者为 0 或者为 1。

$$\gamma = \begin{cases} 0 & \text{当}(1 - p_1(e_1))p_2^f(e_2^f)[\upsilon Q_2^f - \omega_2^f] \leqslant p_1(e_1)U_2^f \\ 1 & \end{cases}$$

此处 e_1、e_2^f、ω_2^f、U_2^f的取值是当 ω_1、ω_2^s、ω_2^f的取值在约束条件 $\gamma = 1$ 下，最大化 π 时的值。[29] $\gamma = 1$ 成立的充要条件是之前项目失败了的工人能够带给公司的期望利润必须要大于第一期工资的节省（保持第一期的努力程度不变），公司可以通过解雇项目失败了的工人而受益，而不是通过用更低的工资合同来惩罚他们。公司获得的好处为 U_2^f的一定比例，即乘以第一期工资被支付的概率 $p_1(e_1)$，而 U_2^f为工人如果在第二期重新被雇佣时获得的效用。需要注意的是劳动力市场上的该结论相当于信贷市场上违约者获得续借贷款所需满足的条件。

求出 U 对 e_2^s 和 e_2^f的微分，我们发现 e_2^s 只是 ω_2^s 的函数，而 e_2^f只是 ω_2^f的函数，用 $S_i(e_i) = p_i(e_i)\upsilon Q_i - e_i$ 表示第 i 期对努力程度 e_i 的社会回报，用 $\pi_i = p_i(e_i)[\upsilon Q_i - \omega_i]$ 表示银行在第 i 期通过支付 ω_i 的工资和诱导出 e_i 的努力所获得的期望利润。那么，我们会发现当效用约束条件没起到约束作用时，公司选择满足下式的 ω_2^s 和 ω_2^f：

$$\frac{d\pi}{d\omega_2^s} = p_1(e_1^s)S'_2(e_2^s)\frac{de_2^s}{d\omega_2^s} = 0\text{，或者 } S'_2(e_2^s) = 0\text{，}$$

$$\frac{d\pi}{d\omega_2^f} = \gamma\frac{d\pi_2}{d\omega_2^f} - p_1(e_1)\upsilon Q_2\frac{\partial p_2^f(e_2^f)}{\partial e_2^f}\frac{de_2^f}{d\omega_2^f} = 0$$

这些一阶条件说明如果第一期的项目成功了，选择的第二期合同能够引发社会最优的努力程度。注意这个结论类似于借贷市场中的结论——第一期项目成功了的借款人将会选择社会最优的投资。如果第一期项目失败了，且 π_2 为 ω_2^f的凹函数，则提供的工资将低于能最大化公司第二期利润的工资水平$\hat{\omega}_2$。当然，如果我们加入一个可实施性条件，工资水平将为$\hat{\omega}_2$。直观的解释是公司通过其对工人前一期努力程度的影响来惩罚表现差的人，从而获得额外的利润。即使加入了可实施条件，也有 $\omega_2^s > \omega_2^f$：用更好的合同来奖励有好历史记录的人。

将分析扩展到完全市场均衡时，第三部分从信贷市场上推导出来的结论仍然成立，只是形式做适当的修改。市场均衡可能是 γ 取任意 0 到 1 之

29 $\partial\pi/\partial\gamma$ 不能在 $\gamma = 0$ 的条件下计算；当 $\gamma = 0$ 时，ω_2^f 的任意值都满足利润最大化。

间的值，所以均衡可能是一些、所有或者没有表现差的工人被解雇，并且 $\omega_2^s > \omega_2^f$，即留下的表现差的工人获得的合同要比留下的表现好的人获得的合同差。

七、结论

我们已经证明了在存在严重的道德风险问题的市场中，一般而言，竞争行为具有跨期联系（intertemporal linkage）的特点。[30]在这些市场中的合同可能包括意外条款，规定了在一定情况下委托—代理关系会被终止，即使委托人为获得代理人而竞争，而所有被终止合同的代理人的回报方程都随机占优于新代理人的回报方程。也就是有经验的借款人被拒绝贷款，工人被解雇，佃农被收回租地，即使在每种情况下，被终止了的借款人、工人或者佃农的期望表现，都会好于他们的替代者。此外，带意外条款的合同的每一条款都会被委托人或者代理人自愿地实施。而且，即使被终止的代理人并不比留下的人更差，但没有别的委托人会向他们提供合同了。

虽然本文关注的是通过拒绝向违约者提供贷款或者解雇雇员的方式，合同之间产生跨期联系的情况，但我们也观察到当价格（租金、工资）视之前的行为而定时，也会产生这种联系。当以下情况之一出现时，需要以终止委托—代理关系的方式建立联系：

（a）使代理人受损的合同条款也会使委托人受损。例如，如果在更低的工资水平下，工人工作得更少，或者如果当工资被降低时，好工人更倾向于退出，那么降低产出低的工人的工资不如解雇这些工人。

（b）最低工资法或高利贷法等法律限制了酬金。

（c）在前期回报低的（失败的）代理人更有可能在后期获得低的回报。

如我们的模型所示，如果有经验的但被终止的代理人比新雇佣的代理人更有能力，那么禁止终止行为的政府政策将使每个人的福利增加。在我们1980年的论文中，我们构建了一系列有该特性的政府干预措施。我们证明了虽然市场均衡是帕累托无效的，但是没有银行或银行联盟可以实现帕累托

30 我们此处用的“竞争性”是指企业都比较小，没有垄断权，而不是指均衡——即没有代理人可以收取市场出清价格以外的价格。在那种情况下，显而易见，如果均衡存在，则为市场出清。在我们的模型中，存在市场出清的价格（利率），但我们已经证明了如果那些利率太高，任何银行都可以通过降低利率来增加它给每个借款人带来的利润。因为任何银行都有动力偏离市场出清的价格（利率），所以说，这些价格是由竞争的力量决定是不太合理的。

改进。[31]

我们关于终止的激励效应的结论是从假设所有的代理人同质的模型中得到的。如果我们将代理人之间的不同处考虑进来，会加强我们的论述。某人的贷款申请被拒绝（某人被解雇）可以传递关于某人作为借贷人（雇员）的受欢迎程度的信息。如布鲁斯·格林瓦尔德（Bruce Greenwald）所证明的，我们可以将乔治·阿克洛夫（George Akerlof）关于二手车市场萎缩或者不存在的讨论扩展到"二手的跨期劳动力市场"或者"二手借款人市场"。

合同之间的联系说明事前和事后竞争的重要区别：在合同被签之前，银行（雇主）为了客户（工人）而竞争，签了合同以后，感到不满意的借款人（工人）不能低成本地转向其他银行（雇主）。只存在有限的事后竞争。这种有限竞争对合同和服务的跨期联系很重要，其确保了长期的承诺。以后更进一步的研究将关注有限竞争在一个更复杂的模型中的作用，该模型将考虑存在异质的委托人和代理人时的市场，且参与人都不完全了解其他参与人的特征。

参考文献

Akerlof, *George*, "*The Market for* '*Lemons*': *Qualitative Uncertainty and the Market Mechanism*," Quarterly Journal of Economics, *August* 1970, 89, 488 – 500.

Allen, *Franklin*, "*Control of Capital*, *Sharecropping*, *Ability and Information*," *mimeo.*, *Nuffield College*, *Oxford*, 1981.

Benjamin, *Daniel*, "*The Use of Collateral to Enforce Debt Contracts*," Economic Inquiry, *July* 1978, 16, 333 – 359.

Calvo, *Guillermo A. and Wellisz*, *Stanislaw*, "*Supervision*, *Loss of Control*, *and the Optimum Size of the Firm*," Journal of Political Economy, *October* 1978, 86, 943 – 952.

Greenwald, *Bruce*, The Labor Market as a Market for Lemons, *New York*: *Garland Publishing*, 1979.

Guasch, *J. Luis and Weiss*, *Andrew*, "*An Equilibrium Analysis of Wage-Productivity Gaps*," Review of Economic Studies, *October* 1982, 49, 485 – 497.

Keeton, *William R.*, Equilibrium Credit Rationing, *New York*: *Garland Publishing*, 1979.

Lazear, *Edward*, "*Why is There Mandatory Retirement*?," Journal of Political Economy, *December* 1979, 87, 1261 – 1284.

---

31 我们1980年论文中所给出的干预的顺序是：先是高利贷法，然后是对贷款给新的借款人的比重的暂时性限制，最后是禁止银行拒绝将贷款给违约者。每个干预都伴随着一定的税收和转移政策以确保借款人和存款人都不会受损。有必要对干预排序，以确保没有任何一群借款人或存款人在向新的稳态均衡转移过程中受到损失。

Ordover, Janusz and Weiss, Andrew, "Information and the Law: Evaluating Legal Restrictions on Competitive Contracts," American Economic Review Proceedings, *May* 1981, 71, 399-404.

Shapiro, Carl and Stiglitz, Joseph E., "Equilibrium Unemployment as a Worker Discipline Device," unpublished paper, 1981.

Stiglitz, Joseph E., "Prices and Queues as Screening Devices in Competitive Economics," IMSSS Technical Report, Stanford University, 1976.

______, *"Incentives and Risk Sharing in Sharecropping,"* Review of Economic Studies, *April* 1974, 41, 219-255.

______ *and Weiss, Andrew, "Credit Rationing in Markets with Imperfect Information, Part II: Constraints as Incentive Devices," Bell Laboratories Discussion Paper No.* 180, *November* 1980.

______ *and* ______, *"Credit Rationing in Markets with Imperfect Information,"* American Economic Review, *June* 1981, 71, 393-411.

Weiss, Andrew, "Job Queues in Labor Markets With Flexible Wages," Journal of Political Economy, *May* 1980, 88, 526-538.

Wette, Hildegard, "Collateral in Credit Rationing in Markets with Imperfect Information: Note," American Economic Revew, *June* 1983, 73, 442-445.

同业监督与信贷市场*

机构性贷款人面临的一个主要问题是如何确保借款人审慎地使用资金，以便提高贷款的偿还可能性。部分解决问题的一个方法是“同业监督”：要求其他人监督借款人，如果借款人破产，则监督人需要支付一定的罚金。孟加拉国乡村银行（Grameen Bank）以及其他类似的团体贷款项目的金融绩效，主要归功于同业监督。但同业监督是有成本的。它将风险从银行转移到了合同的联合签署人上。本文用在竞争性信贷市场上的同业监督的简单模型，证明这种风险的转移导致借款人福利的增加。

获得资金的困难和为所得资金支付的高成本或许是妨碍生产力提高的一个重要因素。农村的资本市场常常是不发达的。农村有传统的放债人，但他们往往因为收取高利贷利息而遭人唾骂。这种高利率存在的原因是一个具有争议的话题。流行的观点认为高利率是剥削性的，这种观点暗含的假设是竞争是有限的。当地的放债人利用了当地信息（local knowledge），而这种当地信息或许就解释了为什么竞争是有限的。最新的观点对“剥削性利率”的说法提出质疑，认为高利率是三个因素的产物：高的违约率、违约事件之间高的相关性，以及甄别贷款申请和发现有违约记录的借款人所需要的高成本。由于当地信息的重要性，放债人的贷款往往集中于某一个地区以内，多样化的缺乏意味着贷款人必须承担很大的风险。

在利率的收取和贷款业务扩展的制度性安排方面，传统的放债人和在发达经济体中建立起来的现代银行机构都有显著的不同。因此，许多政府都鼓励正式的银行机构进入农村。这些机构既可以增加经济效率（通过更广泛

---

* “Peer Monitoring and Credit Markets”, *World Bank Economic Review*, (3), September 1990, pp. 351 – 366. 这一章的内容是建立在早期关于道德风险一般理论的研究（Arnott and Stiglitz, 1985, 1986, 1988）以及和理查德·阿诺特（Richard Arnott）对于同业监督一般理论研究（Arnott and Stiglitz, 1991）的基础上的。本文还大量运用了斯蒂格利茨早期和安德鲁·魏斯（Andrew Weiss）所做的关于信贷市场理论的研究成果（Stiglitz and Weiss, 1981, 1983, 1986, 1987a, 1987b）。非常感谢国家科学基金会、俄林基金会、胡佛研究所的资助。我非常感谢 Richard Arnott 和 Karla Hoff，他们提供了非常有益的意见。

地贷款实现），又可以增加公平性（通过降低贫困农民需要支付的利率实现）。不管高利率反映了由于有限竞争导致的剥削，还是反映了对承受了单一风险的当地放债人的补偿，相信以上观点应该是成立的。由此推断，更有效率的现代机构将把效率低下的当地放债人驱逐出市场。

然而，事实上传统的放债人和现代放贷机构不仅共存，而且当地放债人似乎还能够继续以较高的利率放债。虽然正式的贷款机构常常承受了很大的损失，但当地放债人不仅存活了下来，而且在一些地方还很兴旺。当地放债人能够存活的部分原因是，正式机构没有使所有希望从它们那里贷款的农民获得贷款（或者并没有提供给他们想要的足够多的贷款）。但另一部分原因可能是当地放债人相对于正式机构有一个重要的优势：他们掌握着关于借款人更详细的信息。因此他们能够区分高风险和低风险的借款人，收取合适的利率，并且他们可以更有效地监督借款人，确保资金被有效地使用，从而能降低违约的概率。[1]

在农村地区开展借贷业务的发展中国家的银行中，孟加拉国的乡村银行似乎是一个成功的典型。它向没有土地的或只有很少土地的农民提供小额贷款——平均额度为 70 美元左右。该银行每个月大约要放出 475 000 笔贷款，违约率为 2% 左右，这同其他贷款人 60% ~70% 的违约率形成了很大的反差（Lurie，1988）。乡村银行有许多与众不同的特点，但我在这里想要强调的一点是这些贷款都贷给由 5 个住在同一个村庄的村民组成的小组，这 5 个人要一起为偿还贷款而负责。小组中任何成员不遵守银行的规定，该银行就会剥夺小组其他成员贷款的机会。[2]

因此，乡村银行能够挖掘出关于小组成员的当地信息。它设计的激励机制使得村庄里的其他成员做同业监督工作。此外，也有学者认为在许多信息成本很高的情况下，同业监督或许是一个有效的办法，可以用其设计激励—监督体系（Arnott and stiglitz，1991）。[3]

同业监督并不是没有成本的。乡村银行的借款小组的成员承担了风险，而该风险在不存在监督问题的时候能够更好地被银行所吸收。事实上，借款小组这种形式是人为地在小组成员之间制造了相互依赖关系，导致他们各自

---

1 激励（道德风险）和选择问题是任何信贷市场都会面临的两个核心问题。

2 乡村银行用个人贷款和不同规模的小组贷款做试验。通过反复试验，该银行定下了 5 人规模的小组贷款。“来自同一个村庄的空间上和社会关系上的紧密联系（也）成为了运作良好、有凝聚性小组的主要基础。”（Fuglesang and Chandler，1988，p. 56）

3 例如在劳动力市场上，工人往往比经理更了解自己的同伴是否怠工了。在保险市场上，家庭成员要比保险公司更了解对各种投保了的事件应该采取怎样的防范措施。

承担了更多的风险。

这就引起了一个分析上的问题：提高监督带来的好处是否大于增加了相互依赖所带来的成本？这个问题就是本文所要建立模型分析并解答的。本文可被视作是为了建立一个关于同业监督的一般性理论而做的第一次尝试。[4] 因此假设借款小组由两个个体组成，而且他们之间的相互依赖是有限的——在同伴违约时，另一人只需要支付有限量的罚金。但是即使是有限量的罚金也使他们承担的风险增加了。同时我还假设小组成员拥有的关于另一成员的信息本身是无成本的，它是相互邻近居住的副产品（在更一般的情况下，监督的程度取决于相互依赖的程度，所以如果相互依赖的程度很低，相互之间的监督也就是有限的）。最后，我假设违约的风险是独立的。在实际中，它们是有关联的。相互关联的存在只会加强分析的结论。

下一部分给出了基本的模型，该模型描述了缺乏同业监督时的均衡。接下来的一部分会证明同业监督是如何起作用的，并解释为什么它会被采用。

一、基本模型

我假设所有个体都有两个项目可以实施：一个相对安全的项目，如果以 L 的规模（以花费的美元衡量规模）实施并成功了，会带来 $Y_S(L)$ 的回报；另一个相对风险较高的项目，如果成功了会带来 $Y_R(L)$ 的回报。如果项目失败了，回报为0。项目成功的概率分别为 p_S 和 p_R，有 $p_S > p_R$。我假设回报是规模的增函数，风险高的项目的固定成本，$\bar{L}$，要比安全的项目的固定成本高：$\bar{L}_R > \bar{L}_S$。因此，在相应的区域，Y'_R 要大于 Y'_S，如图1所示。

考虑到成功的概率，假设安全的项目总是比高风险项目带来更高的回报：

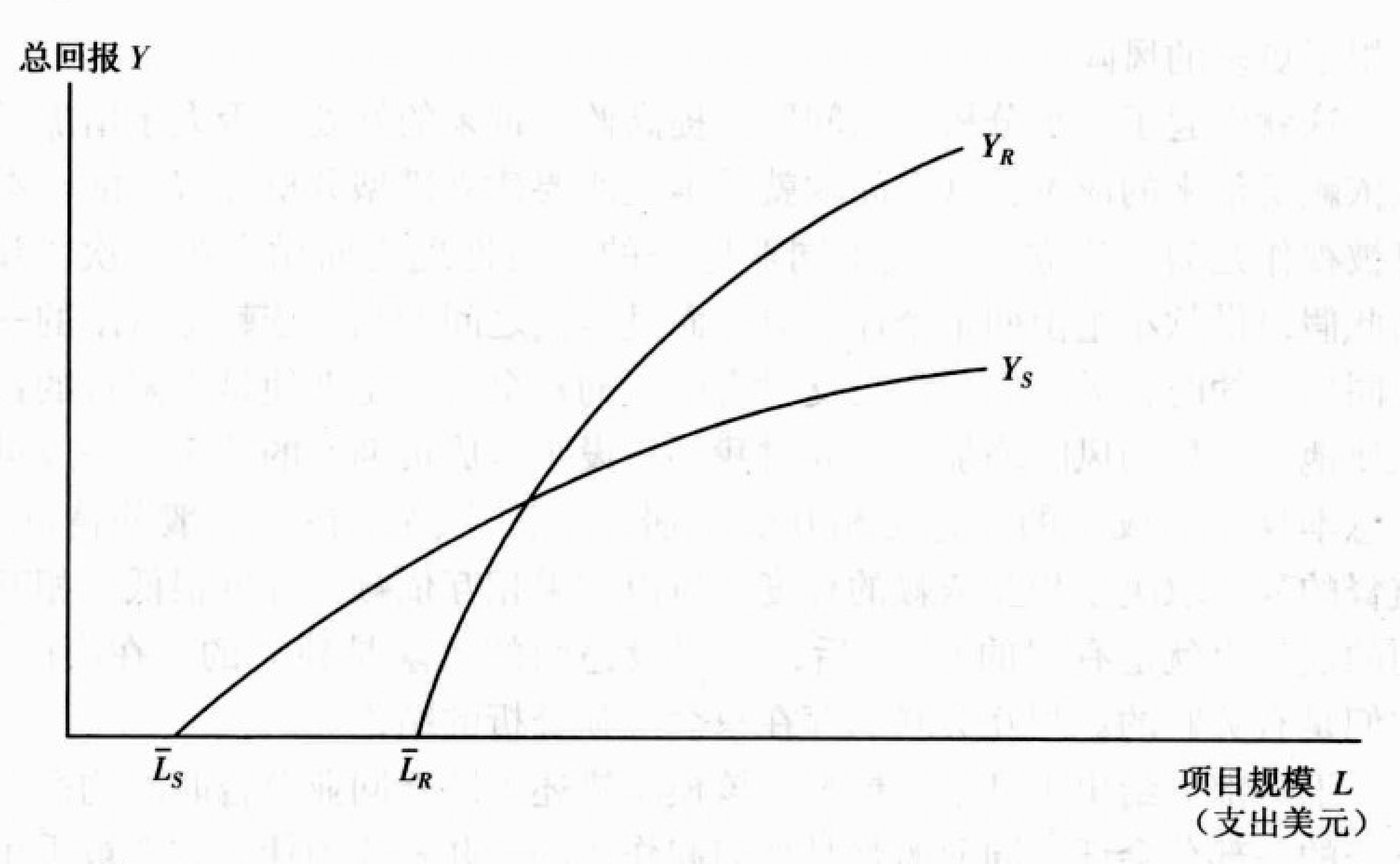
$$Y_S(L)p_S - [1+r]L > Y_R(L)p_R - [1+r]L \quad \forall L \tag{1}$$

r 是利率。因此投入自己资金的个体一直会选择安全的项目。然而，用借来的资金进行投资，并在项目失败后宣告破产的个体，会考虑到破产的概率而把资金的成本打个折扣。

为了集中讨论激励问题，我假设所有的个体都是一样的，并且为了简化

---

4 该模型关注的是小组贷款这种形式是如何影响到借款人对项目的选择以及他们的偿还能力的，而不是关注如何迫使有能力偿还贷款的借款人偿还贷款的问题。更进一步的研究（Besley and Coate，1991）讨论了小组贷款在解决后一个问题中起到的作用。读者还可以参见 Varian（1990）。



注：$\bar{L}$ = 固定成本；R = 风险项目；S = 安全的项目

图1　安全的高风险项目总回报和投资（假设成功）之间的关系

分析，假设在任一规模上，两个项目所要求投入的努力程度是一样的。[5]实施项目 i 带来的期望效用为：[6]

$$V_i(L,r) = U(Y_i(L) - [1+r]L)p_i - \nu(e(L)) \tag{2}$$

此处 $U(Y)$ 为收入的效用，$U'>0$，$U''<0$，效用函数被标准化，使得 $U(0)=0$。[7]$\nu(e(L))$ 是投入努力 e 带来的负效用，有 $\nu'>0$，$\nu''>0$。假设当项目规模增加时，所要求投入的努力程度也上升：$e'(L)>0$。图2给出了个人对某一给定项目（高风险或安全的）的无差异曲线。该曲线给出了所有使借款人

5　如果村民之间互相非常了解，那么在形成同业监督小组的过程中，如果没有社会力量起到反向作用，将会出现"同型配对"（assortative mating）的现象，即最不可能违约的人组合在一起，违约可能性大一些的人组合在一起，如此，等等，最后使得违约可能性最大的人形成一个小组。因此认为一个同业监督小组中的成员都是一样的假设，可以被视作是均衡条件之一，我们可以在更一般的条件下推导出这个条件。

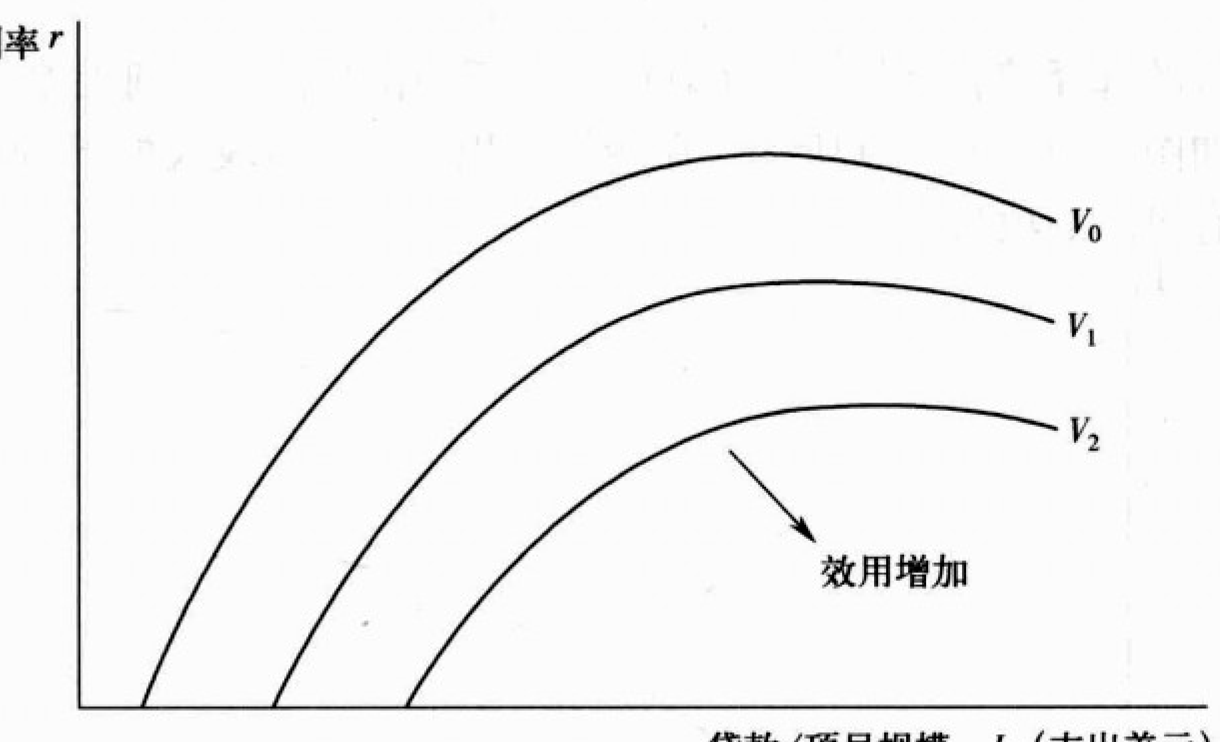
当小组的形成由全社会来决定时，同业监督可能成为一个比本文的分析所建议的更能解决道德风险问题的办法。

6　我假设借款人除了从项目中获得收入以外没有其他的收入来源；或者有收入，但该收入是常数并且在项目失败的时候不能被银行扣押。

7　这种标准化使本文的表达更方便，但并不重要，而且可能会在效用函数的绝对风险厌恶系数为常数的情况下遇到麻烦。

这个公式的含义是借款人在项目上的投资等于他能够借到的资金数额 L。这个结论可以被一般化，也适用于借款人愿意投入的自有资金的数额依赖于他可以借到的数额的情况。

获得相同效用的合同(L, r)的组合。[8]



注：V_i 为第 i 个项目的期望效用，其中 $V_0 < V_1 < V_2$。

图 2　某一项目贷款规模和率利的无差异曲线

如果借款人选择实施项目 i，则无差异曲线的斜率为：[9]

$$\frac{dr}{dL} = \frac{Y'_i - [1 + r] - \nu' e' / U' p_i}{L} \tag{3}$$

"转换线"（switch line）可以被定义为使个体在两个项目间无差异的(L, r)的组合，也就是满足：

$$V_S(L, r) = V_R(L, r) \tag{4}$$

因为规模报酬对高风险项目而言比对安全项目更重要，所以保持 r 不变，L 的增加使高风险项目更有吸引力，在这种可能的情形下，转换线的斜率为负。在相应区域，即当 $L > L_R$ 时，我们有：

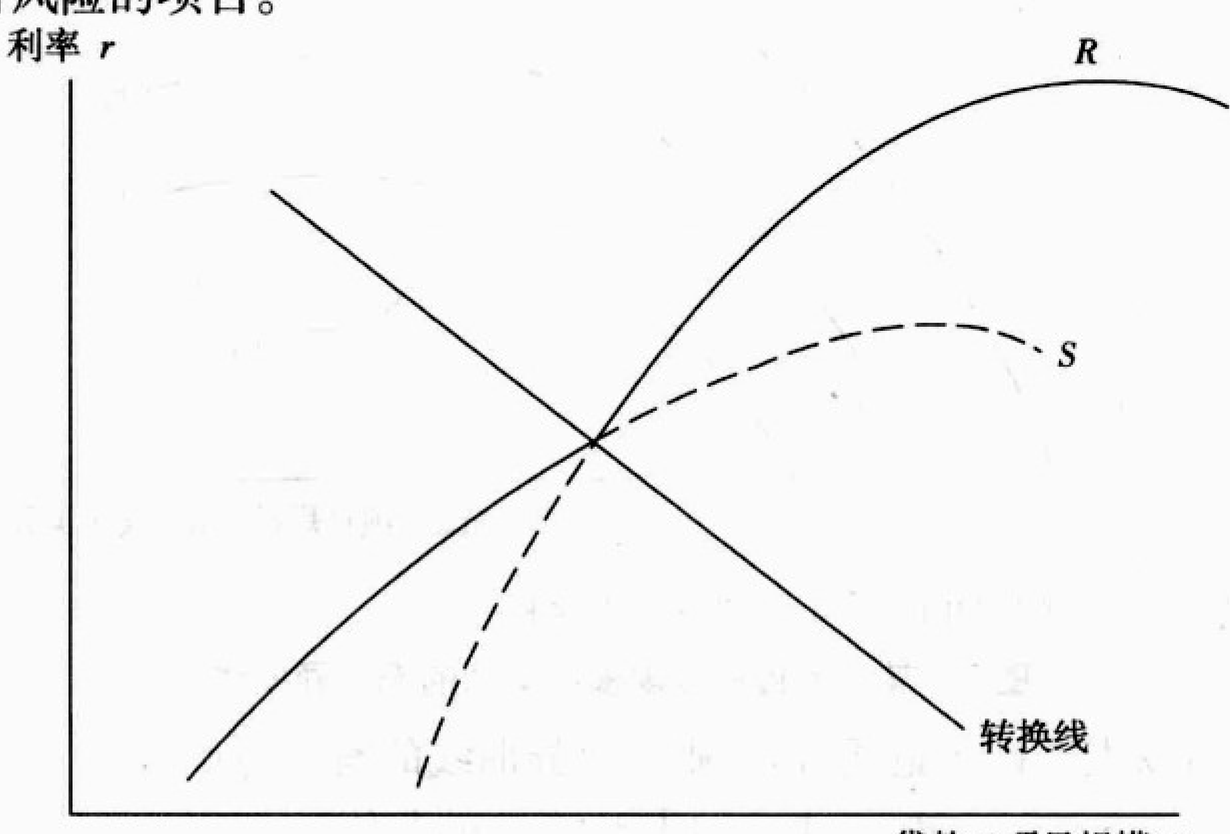
8　我假设虽然贷款人不能监督借款人的其他行为，但贷款人可以监督借款人的借贷行为，确保他没有获得资金的其他来源。这个假设不能被完全满足。贷款人可以限制他提供的贷款的数额，但正规的贷款人常常难以限制其他贷款人提供的贷款。因此本书第一部分中的几个案例分析都说明了当信息和其他的交易成本使得借款人只同一个（或者，最多，很少几个）非正式贷款人有借贷关系时，借款人常常会同时从正式和非正式的信贷机构借钱。

完整地分析当正式机构不能限制外来贷款的数额时的市场均衡已经超出了本章的范围（参见 Arnott and Stiglitz，1991，所做的对存在道德风险的保险市场上的类似问题的分析）。然而这样的分析可以进一步说明同业监督的作用，因为当没有同业监督时，无法限制外来贷款将降低借款人在正式的信贷市场上的期望效用水平。

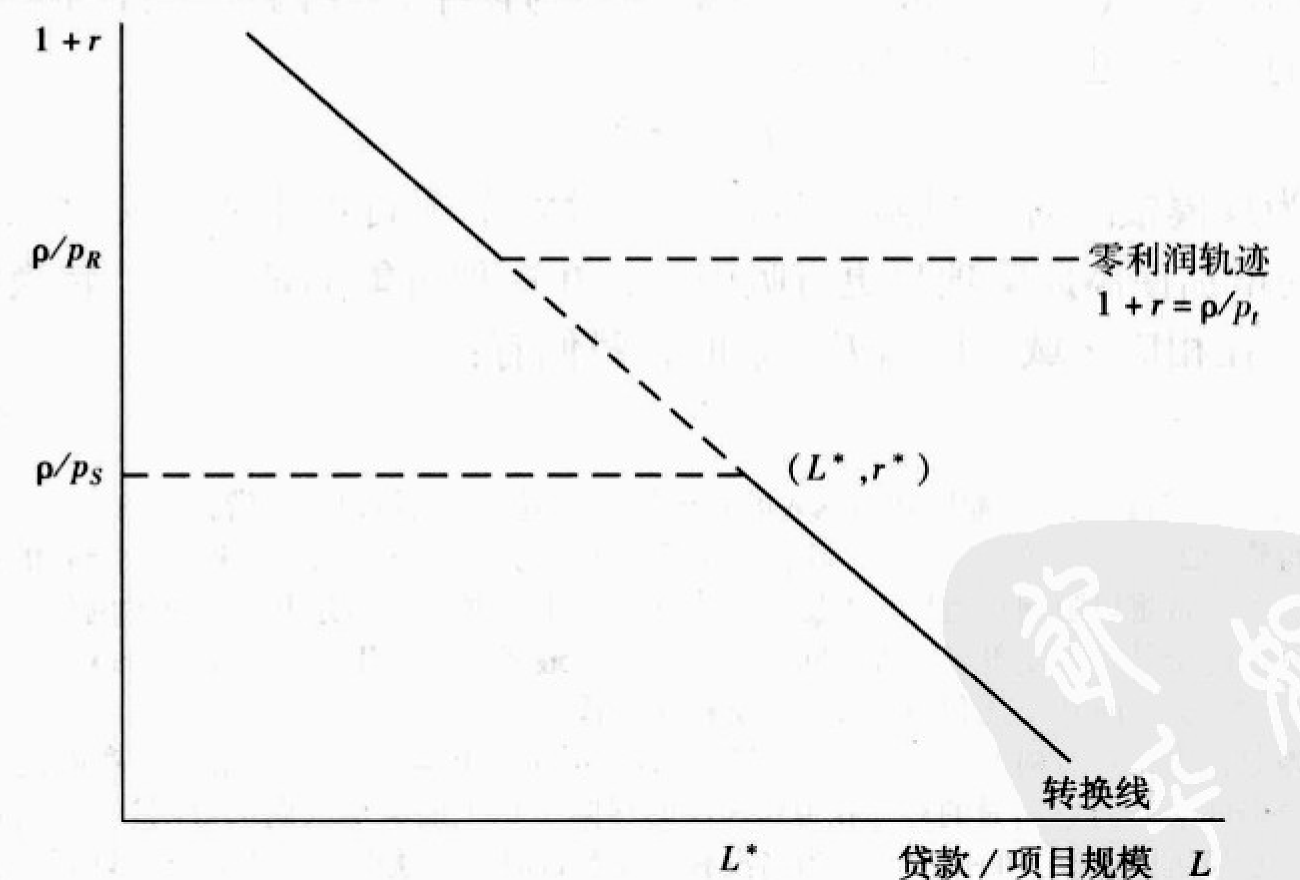
9　某一给定项目的无差异曲线在一定区域内是"表现良好的"——若 $Y''_i < 0$，$d^2(\nu' e')/dL^2 > 0$，则 $Y'_i > [1 + r]$。

$$\frac{\partial V_S}{\partial L} < \frac{\partial V_R}{\partial L} \tag{5}$$

当对项目的选择随着贷款合同条款的变化而变化时，注意到无差异曲线为扇形的，如图3的上面部分所示。在转换线以上（L 值较大的部分），借款人将实施有风险的项目。



注：R = 高风险项目；S = 项目。由于大规模贷款可使贷款人承受有风险的项目，无差异曲线（使各借贷合同采用的技术不同）产生扇形斜率。



注：当合同为（L^*，r^*），出现市场均衡，此时利润为零。这是零利润轨迹上贷款规模最大的点，借款人愿意从事安全项目。参数 ρ = 资金成本；p_i = 项目成功概率 i（$i = R$，S）。

图3 选择安全项目和高风险项目时贷款规模和利率的影响

注意到转换线是向下倾斜的，若贷款规模固定不变，则随着 r 的增加，效用会降低 $LU'p_i$。由于高风险项目的 U' 和 p_i 更低，对每单位 r 的增加，高风险项目的效用的减少要少一些。因此，从使借款人在实施安全项目和高风险项目之间无差异的(L, r)值——如图 4 中的 E 点——开始，r 的增加会使高风险的项目占优于安全的项目。但在式（5）中假设了当 r 固定时，L 的增加使高风险项目带来的期望效用的增加大于使安全项目带来的期望效用的增加。因此 L 的增加必须伴随着 r 的降低，才能使借款人在两个项目之间无差异，这证明了转换线的斜率为负。

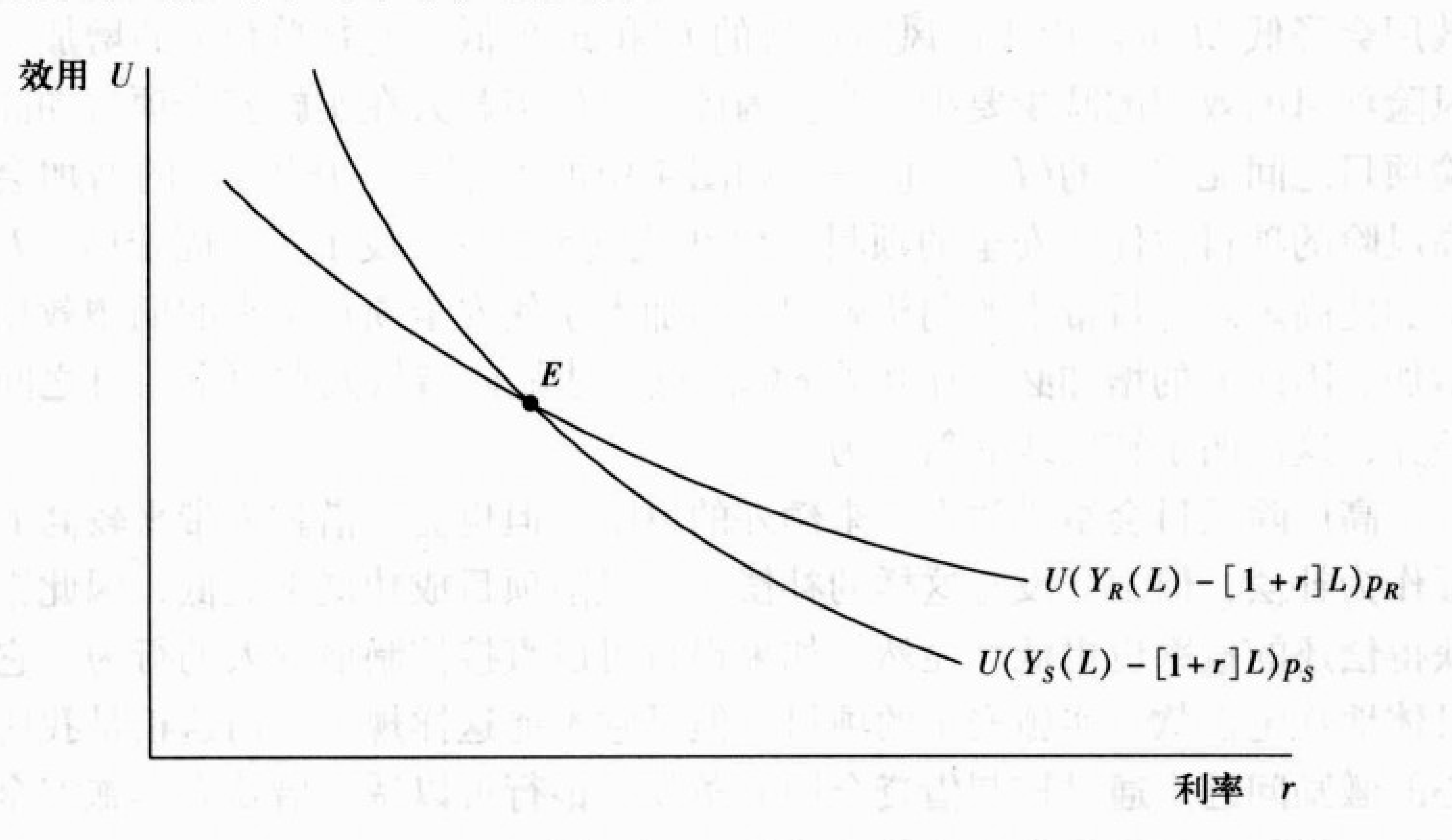
高风险项目会给借款人带来额外的风险，但也会给借款人带来较高的回报作为补偿，但银行没有这样的补偿。高风险项目成功概率更低，因此银行获得偿还的概率也更低。显然，如果银行可以直接控制借款人的行为，它会具体地规定借款人实施安全的项目。但是它不能这样规定，而这正是我所关心的激励问题。通过控制借贷合同的条款，银行可以诱导借款人实施安全的项目。银行必须提供在转换线上或以下的合同。

为了分析市场均衡，还需要有另外一组曲线——零利润轨迹线。用下面的方法可以很容易地构造零利润轨迹。如果借款人实施安全项目，则银行的期望回报为 $p_S[1+r]$。如果资金成本为 ρ，那么如果 $1+r=\rho/p_S$，则利润为零。同样地，如果借款人实施高风险项目，则当 $1+r=\rho/p_R$ 时利润为零。因此零利润轨迹为图 3 下半部分的虚线曲线。

市场均衡是零利润轨迹上最大化了借款人期望效用的点（假设借款人没有其他的贷款途径或者，等价地，贷方可以监督借款人的贷款总额）。在图 3 的下半部分，均衡贷款合同是(L^*, r^*)。显然，借款人在市场利率下愿意借更多的钱，而如果借款人可以可信地承诺不实施高风险的项目，则贷款人也愿意在那个利率下借给他更多的钱。但倘若借款人无法使自己的承诺可信，就算其做出了承诺，但因贷款人无法强制实施该承诺（借款人和贷款人都知道这点），那么贷款人必须将贷款规模限制为 L^*。

这只是在更一般模型中，信贷配给所可能采用的两种形式中的一种。还有一种可能采取的形式是在一组完全一样的借款人中，一些人获得了贷款而另外一些人没有。为什么这种信贷配给形式在这些人中没有发生作用？常见的解释是那些没有被分配到贷款的人会提出支付更高的利息。当他们这样做时，利率会被抬高直到对资金的需求等于供给。从这个模型中，我们很容易就可以看出为什么这个解释是不对的。贷方知道高于转换线的任一利率水平，借方都会实施高风险的项目。虽然借款人承诺支付的回报更高，但其实

际支付的回报（平均水平）会更低。[10]



注：L = 贷款/项目规模；p_i = 项目成功概率 i（$i=R$，S）；Y_i = 如果成功，项目 i 的总回报。

图4 一定贷款规模下选择安全项目和高风险项目时利率对效用的影响

二、同业监督

现在，假设每个借款人都有一个（且只有一个）同样是借款人的邻居。他们各自项目的成功与否是相互独立的。两个借款人可以相互监督。贷款人希望他们各自报告邻居是否实施了高风险的项目。贷款人要创造一种条件，使得每个借款人监督同伴并报告同伴的任何欺骗行为是对其自己有利的。

下面是达到该目的的一种简单的方法。银行提供一种合同，承诺如果借款人的邻居同意联合签署——承担下面描述的责任——则借款人可以获得较低的利率并获得更多的资金。联合签署者则同意，当联合签署的贷款被违约时，他要向贷款人支付 qL 美元的罚金——当然，要在他自己没有违约的情况下。

---

10 在斯蒂格利茨和魏斯（Stiglitz and Weiss，1981）的文章中对这个问题进行了详细的讨论。在这里给出的简单模型中，在给定的利率水平下，贷款人对借出转换线以下的任意规模的贷款都是无差异的。但是如果对模型进行一点修改，允许 p_i 随着贷款规模的增加而增加少许，那么在转换线以下，零利润轨迹的斜率为负，将贷款的规模降低到 L^* 以下，实际会降低贷方的期望回报。

为了使模型更符合现实，对模型所做的另一个修改解释了为什么贷方不愿意提供小额贷款给“资金不足”的项目。例如，借款人往往有办法采取行动，使贷款人不得不贷出更多的钱或者在承担损失之前借出贷款的风险。因此借款人可以“迫使”贷款人借给他们更多的钱。参见 Stiglitz and Weiss（1981）以及 Hellwig（1977）。

现在，联合签署者的期望效用依赖于其邻居实施的项目是高风险的还是安全的。给定我们已在这个问题上所做的独立性和对称性的假设，推断他们将合作是合理的。也就是他们一起决定实施安全项目还是高风险项目，而且如果他们实施了高风险项目，他们会达成一致，不向银行报告。[11]

联合签署邻居的贷款合同，会使借款人承担额外的风险。因为零利润条件确保利率将会调整到使银行的期望回报不变——将联合签署者的支付计算进来——联合签署的规定导致借款人的收入在任一给定贷款规模 L 下都出现了均值保留展型。如果两个借款人都成功了，他们的收入和效用会更高，但如果其中一个成功，另一个失败了，第一个借款人的收入和效用会更低。为了对其承担的额外的风险进行补偿，他必须能够获得更多的贷款。图5 描述了达到一定期望效用水平所要求的最小贷款规模和联合签署者的支付率 q 之间的关系。本文附录

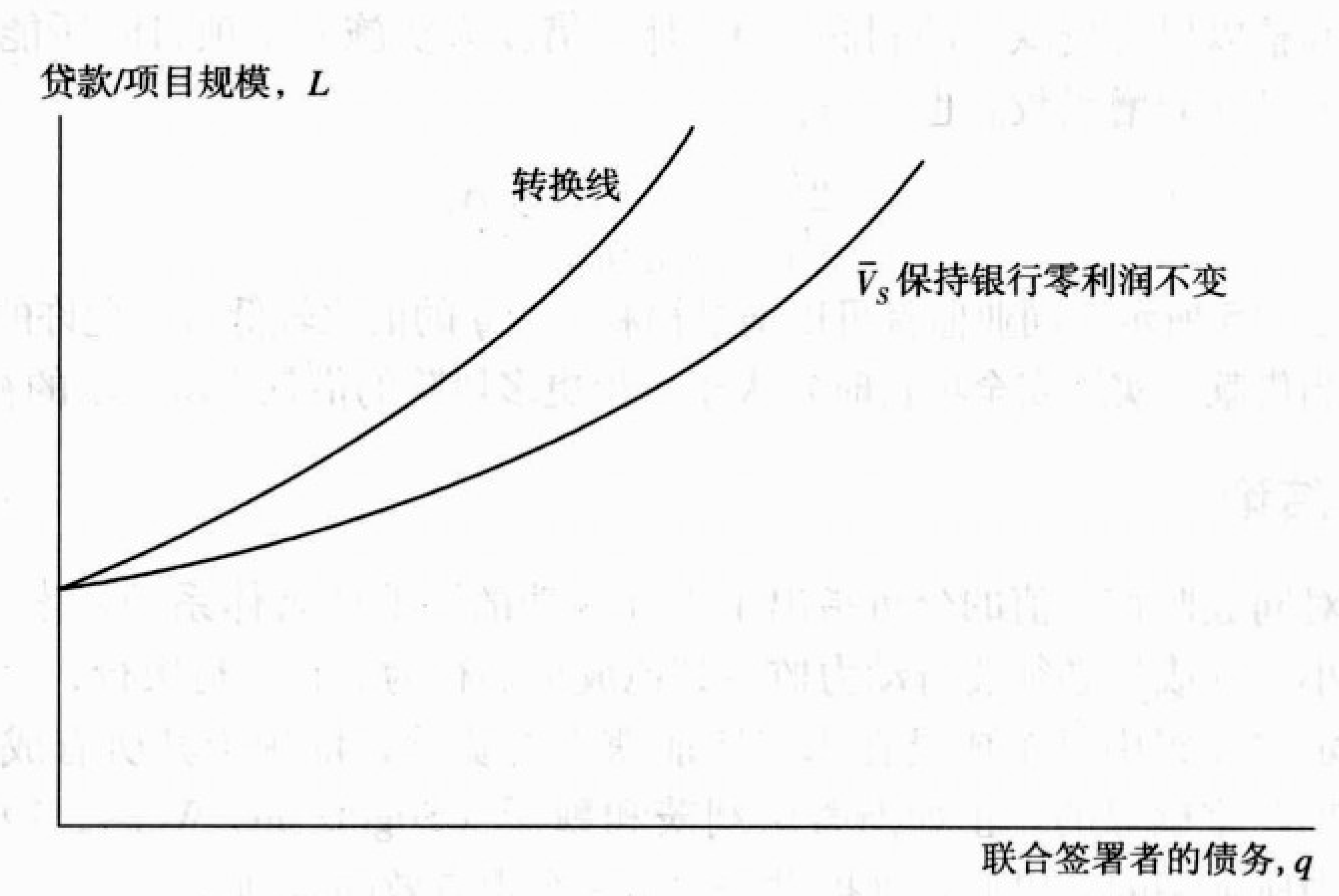


图 5

11 个体间相互作用所实现的均衡，没有在模型里被详细地刻画。构造一个博弈使这成为其均衡结果是比较容易的。例如，假设在任何时候，一方报告他的邻居实施了高风险的项目，则另一方有时间报告相同的信息。因此任何一方违反不报告的协议都不能获得好处。

更一般化地，假设社会制裁会保证：当他们的收入是相互依赖时，他们合作是合理的。自然信息的假设确保他们不能相互欺骗。

我们始终假设如果借款人通过实施高风险项目来违背合同，联合签署者可以“强迫”其撤销该行为。例如，贷款合同可能规定如果那样的欺骗发生，该贷款被违约，贷款人获得该项目的所有回报。

中的式（13）证明了，给定银行的零利润条件，在$q=0$时，我们有：

$$\left.\frac{dL}{dq}\right|_{\bar{V}_S} = 0 \quad (6)$$

这意味着（在q值较低时）联合签署给借款人带来的风险负担正好被所收取的竞争性利率的降低所补偿了。

剩下的问题只是确定转换线有什么变化了。如果两个借款人合作，那么转换线则为：

$$U(Y_S[L] - [1+r]L)p_S^2 + U(Y_S(L) - [1+r-q]L)p_S[1-p_S]$$
$$= U(Y_R[L] - [1+r]L)p_R^2 + U(Y_R(L) - [1+r-q]L)p_R[1-p_R] \quad (7)$$

本文附录中的式（15）说明了只要式（5）的假设满足了并且当q上升时，利率调整以保持贷款人的利润为0，那么借款人实施安全项目时所能贷到的最大L是q的增函数。也就是：

$$\left.\frac{dL}{dq}\right|_{q=0\text{时的转换线}} > 0 \quad (8)$$

如图5所示，同业监督可以增进福利。当q的值比较低时，允许的L的增加（当借款人实施安全项目时）大于承担更多风险的借款人所要求的补偿。

三、结论

对同业监督价值的分析指出了设计成功的同业监督体系的要素。第一，联合小组的成员必须要有动力监督其他成员的行为。在乡村银行，这点是通过规定“小组中每个成员在未来所能获得的贷款，依赖于其所有成员的偿还表现”来确保的。正如斯蒂格利茨和魏斯（Stiglitz and Weiss，1983）的研究中所强调的，以后拒绝提供贷款是一个很有效的激励设计。

乡村银行采取小规模组合的形式，虽然小规模增加了单个成员违约所带来的风险，但是增加了同业监督的动力。后者带来的好处很可能是大于后者带来的损失的。规模大的组合存在搭便车的问题——每个人都偏好于其他人投入精力来监督，并避免由于举报使用借贷资金不当的人所可能招致的憎恶。此外，任何成员的违约给每个成员带来的成本可能足够小，导致监督的动力——即使除去搭便车问题——也很小。

风险相似的人有很大的动力去组成小组。因为小组是合作型的，如果某个人比其他人更倾向于违约，他就会相当于获得补贴。当小组成员都一样时，则没有补贴（事前的概念）。当然，违约风险高的人愿意加入违约风险

低的组。当违约风险最低的人意识到他们组成小组后的共同利益时，就会发生选择性分组（assortative grouping，即同类型的人自动组合在一起——译者）。然后，剩下的人中风险最低的人又会组合在一起，这个过程一直进行到使风险最高的个体被迫组合在一起。村民不仅在监督方面，而且在选择方面也比正式的信贷机构有信息优势。选择的有效性通过消除交叉补贴带来了正面的效果，而交叉补贴在甄别不完全的信贷市场中经常发生。同业监督会增强农村信贷市场的有效性。[12]

传统的观点认为，联合签署的规定是增加一笔贷款有效抵押品的办法。本文提供了另外一种解释。联合签署使签约者有动力去监督同他一起贷款的人的行为，并且改变了借款人自身的行为。联合签署也增加了风险。但在这里分析的这种对称性竞争均衡下，利率会因为监督的提高而调整。我的核心结论是当 q 的值很低时，从同业监督中获得的好处多于所承担风险的增加而导致的期望效用的损失。

在发展中国家，村庄以外的贷方无法对贷款实施监督，这导致了发展有效的资本市场的一个主要障碍。在村庄里，风险是高度相关的，且只有极少的人有财富用于信贷，所以信贷市场是不完全竞争的且存在很高的风险升水。

自然地，在这里就会产生一个疑问：如果同业监督如此有效，为什么私人市场不使用它呢？在发达国家的资本市场中，同业监督是被广泛地使用的。就像上面所提到的，联合签署的规定可能不仅对增加有效抵押品很重要，也对导致选择和同业监督效应很重要。

在发展中国家，非正式的贷款人或许并不需要同业监督，因为他们可以直接监督借款人。而对正式机构而言，发展的同业监督——以及发展其他机制的——一个主要障碍来自于法律体系的缺乏，使得其不能强制实施合同。政府比私人贷款人有优势，这是在发展中国家的一个很重要的不同之处，而这些国家的司法体系要么办事效率低下，要么就根本不起作用。政府或许拥有私人贷款人所没有的权力来强制实施合同。

这说明除了更广泛地提供政府信贷以外，另一项改革措施是：改革法律体系以保障贷款人能收回其贷款。然而，将只针对信贷市场的有效性所要进行的法律改革和更广泛的法律改革分开进行是很难的。而推行广泛的法律体系改革会遇到很多严重的阻碍。虽然法律改革可以促使在私人市场上使用同

12 如果不是根据违约率的不同，对不同的组收取不同的利率，则在不同组之间也可能存在交叉补贴。然而，成功的同业监督降低了小组的违约率，使得这种交叉补贴相对不那么重要。

业监督，但即使缺乏这样的根本性的改革，精心设计的、能利用同业监督机制的政府信贷计划或许也是一个有效的次优政策。

附录：

无同业监督

为了简化，我们用 $\bar{r}\equiv 1+r$，表示银行收取的本金和利息；用 $U_i\equiv U(Y_i(L)-\bar{r}L)$ 表示在项目 i 上成功了的借款人的效用；并有 $i=R, S$。[13]

由前文可知 $V_i(r, L)=U_ip_i$ 为实施了项目 i 的借款人的期望效用，转换线是被配给贷款的借款人所面临的合同条款 (L, r) 的集合，即为

$$V_R = V_S \tag{9}$$

在（5）中我们假设在一定的值域中（$L>\bar{L}_R$），多1美元贷款给高风险项目带来的好处要多于给安全项目带来的好处：

$$U'_R[Y'_R-\bar{r}]p_R=\frac{\partial V_R}{\partial L}>\frac{\partial V_S}{\partial L}=U'_S[Y'_S-\bar{r}]p_S \tag{10}$$

对转换线式（9）求全微分，有：

$$\left.\frac{dr}{dL}\right|_{\text{转换线}}=\frac{\left(\frac{\partial V_R}{\partial L}-\frac{\partial V_S}{\partial L}\right)}{L[U'_Rp_R-U'_Sp_S]}<0$$

此处符号条件由式（10）和 $p_R<p_S$、$U'_R<U'_S$ 所得。因此转换线为向下倾斜的，如图3所示。

同业监督

当有同业监督机制时，借款人实际上面临三种可能的结果：1）他自己的项目和邻居的项目都成功；2）他自己的成功但其邻居的失败了；3）他自己的失败了。这三种状态各自带来的效用分别是：

$$U_i\equiv U(Y_i(L)-\bar{r}L)$$
$$U_{iq}\equiv U(Y_i(L)-\bar{r}L-qL)$$
$$U(0)=0$$

对称均衡——借款人和其邻居都选择一样的项目，R 或者 S——下的期望效用是：

13 在整个附录中，忽略掉了管理项目所需要的努力。把贷款规模的变化导致的努力程度变化的各种影响合并在一起是很显然的。

$$\bar{V} = U_i p_i^2 + U_{iq} p_i [1 - p_i] = V_i(r, L, q) \tag{11}$$

假设均衡为信贷配给，银行选择合同（r，L，q）来确保借款人会选择安全的项目。银行的零利润条件是 p_S ［$1+r$］ $+p_S$ ［$1-p_S$］ $q=\rho$，所以

$$\frac{dr}{dq} = -[1 - p_S] \tag{12}$$

对任意的 r，式（11）和式（12）定义了借款人的贷款限额和保持借贷人期望效用不变，且满足银行零利润条件的共付率之间的关系。这种关系可以用下式描述：

$$\left.\frac{1}{L}\frac{dL}{dq}\right|_{\bar{V}\text{和银行的零利润条件}} = \frac{-U'_i p_i^2 [1 - p_S] + U'_{iq} p_i p_S [1 - p_i]}{U'_i [R'_i - \bar{r}] p_i^2 + U'_{iq} [Y'_i - \bar{r} - q] p_i [1 - p_i]}$$

$$= \frac{-M_i}{\partial V_i / \partial L} \tag{13}$$

$$= 0, \text{如果 } q = 0, \text{且 } p_i = p_S \tag{13'}$$

$$> 0, \text{如果 } q = 0, \text{且 } p_i = p_R \tag{13''}$$

此处

$$M_i \equiv \frac{\partial V_i}{\partial q} + \frac{\partial V_i}{\partial r}\frac{dr}{dq}$$

式（13）给出的结论是：在借款人实施安全项目，银行利润为零的均衡下，贷款限额 L 固定时，规定一个低值的联合签署责任率 q，能保持借款人的效用不变。参见图 5 中较低的曲线。

明确地写出转换线式（9）或式（7）有助于分析：

$$p_R^2 U_R + p_R [1 - p_R] U_{Rq} = p_S^2 U_S + p_S [1 - p_S] U_{Sq} \tag{14}$$

对式（14）做全微分得：

$$\left.\frac{dL}{dq}\right|_{\text{转换线}} = -\frac{M_R - M_S}{\partial V_R / \partial L - \partial V_S / \partial L}$$

由式（10）的假设可得，分母是正的。分别用式（13′）和式（13″），我们得到：在 $q=0$ 时，$M_S=0$ 和 $M_R<0$。因此，

$$\left.\frac{dL}{dq}\right|_{\text{转换线}} = \frac{-M_R}{\partial V_R / \partial L - \partial V_S / \partial L} > 0 \tag{15}$$

式（15）说明同业监督使转换线上移。其放松了为了确保借款人实施安全项目而对(L，r)要求的约束。比较式（13′）和式（15），可以看出当 q 值很低时，转换线的上移超过了为了保持借款人的期望效用不变所需要的移

动，如图 5 所示。同业监督因此而增加了借款人的福利。

参考文献

Arnott, Richard, and Joseph E. Stiglitz, 1985, "Labor Turnover, Wage Structure and Moral Hazard: The Inefficiency of Competitive Markets," *Journal of Labor Economics* 3 (4): 434 - 462.

——, 1986, "Moral Hazard and Optimal Commodity Taxation" *Journal of Purblic Economics* 29: 1 - 24.

——, 1988, "The Basic Analytics of Moral Hazard". *Scandinavian Journal of Economics* 90 (3): 383 - 413.

——, 1991, "Moral Hazard and Nonmarket Institutions: Dysfunctional Crowding Out or Peer Monitoring?", *American Economics Review* 81 (1): 179 - 190.

Besley, Timothy, and Stephen Coate, 1991, "Group Lending, Repaytment Incentives and Social Collateral", R. P. D. S. Discussion Paper 152, Woodrow Wilson School, Princeton University, Princeton, N. J.

Fuglesang, Andreas, and Dale Chandler, 1988, *Participation as Process—What We Can Learn from Grameen Bank*, *Bangladesh*, Dhaka: Pearl Printing and Packaging.

Hellwig, Martin F. 1977, "A Model of Borrowing and Lending with Bankruptcy", *Econometrica* 45 (8): 1876 - 1906.

Lurie, T., 1988, "Increasing Income and Employment in Bangladesh", *Ford Foundation Letter* 19 (4): 1 - 5.

Sanderaine, Nimal, 1978, "An Analytical Approach to Small Farmer Loan Defaults", *Savings and Development* 2 (4): 290 - 304.

Stiglitz, Joseph E., and Andrew Weiss, 1981, "Credit Rationing in Markets with Imperfect Information", *American Economic Review* 71 (3): 393 - 410.

——, 1983, "Incentive Effects of Termination: Applications to the Credit and Labor Markets", *American Economic Review* 72 (December): 912 - 927.

——, 1986, "Credit Rationing and Collateral", In Jeremy Edwards, Julian Franks, Colin Mayer, and Stephen Schaefer, eds., *Recent Developments in Corporate Finance*, New York: Cambridge University Press.

——, 1987a, "Credit Rationing with Many Borrowers", *American Economic Review* 77 (March): 228 - 231.

——, 1987b, "Macro-economic Equilibrium and Credit Rationing", National Bureau of Economic Research Working Paper 2164. Cambridge, Mass.

Varian, Hal, 1990, "Monitoring Agents with Other Agents." *Journal of Institutional and Theoretical Economics* 146: 153 - 174.

激励理论

激励、风险与信息：对等级理论的评注*

这篇论文分析了激励、风险和信息在决定雇佣合同结构方面的作用。我们尤其关注比较计件工资和计时工资薪酬体系的作用，以及监督者的作用。对计件工资或计时工资的选择与风险分担有关，与是否将薪酬体系作为甄别雇员的方法有关，也与关于执行任务难度的信息差异问题有关。因此对薪酬体系的选择依赖于雇员和雇主对风险的态度、努力水平的供给弹性、不确定性的来源和大小以及在雇佣关系中所用的监督的特性。监督人的作用被视作是监督投入（执行合同）、甄别工人、获得关于社会状态的信息。而这些作用都同信息的非凸性有关。

一、引言

工人们根据各种合同向企业提供服务。然而，经济学理论却几乎没有对为什么特定的企业要选择特定合同这一现象进行解释。“在竞争性市场上，

* Incentives, Risk, and Information: Notes Towards a Theory of Herarchy, *The Bell Journal of Economics*, Vol. 6, No. 2 (Autumn, 1975), pp. 552 – 579. 本文为修改稿，最初的论文发表于 1973 年的柏林计划研讨会（Berlin Symposium on Planning）。修改过的版本于 1974 年 9 月发表于在宾夕法尼亚大学（University of Pennsylvania）举行的内部组织经济学研讨会（Symposium on the Economics of Internal Organization）上。该研讨会由国家科学基金会（NSF-GS-35889X）和通用电气基金会（General Electric Foundation）资助。作者感谢 Lars Osberg, A. B. Atkinson, A. M. Spence 以及研讨会的与会者，感谢他们所做的评论。作者尤其要感谢 John Pencavel 提出的深刻见解与评论。“An Essay on the Economics of Work Effort and Wage Payments Systems”（Stanford, mimeo, 1974）中的结论和这里的讨论很接近。本研究由斯坦福大学 Institute for Mathematical Studies in the Social Science 的国家科学基金会课题 SOC74 – 22182 资助。

工人会获得等于他们的边际产出的报酬”的断言使人们忽视了这一“细节”。如果存在不确定性，工人将根据所处的情况，做出出售劳动力的决定。

本文的目的是：分析是什么因素决定了特定企业选择特定结构的合同。既然合同在很多条款上都不同——保障薪水的数额、奖金的数额、调整奖金的规则、监督的程度、晋升政策、退休政策等，个人和生产过程的大量细节特点决定了被选用合同的具体结构。因此该领域的理论工作必定会有些烦琐。我们主要将分析集中在工作合同的两个方面：对计件工资和计时工资薪酬制的选择，以及监督的作用。事实上这两方面是紧密相连的：前者可以被视作当某种监督成本非常高时，对均衡工资合同的分析。即使在对工作合同的这两方面所做的分析中，我们也应该将讨论的范围缩小，只讨论在竞争性的劳动力市场上，决定均衡工资合同的一些主要因素。

如论文的题目所示，我们关注的主要是同信息、风险和激励相关的问题。事实上，风险和激励的问题也可以视作不完全信息经济学的一部分：风险反映的是自然状态（例如，降雨量）、个人能力、分配给个人的任务等信息的不完全。传统的风险共担安排在共担风险以后面临可能会降低激励（例如，在保险学文献中的道德风险）的问题，而激励之所以成为一个问题，是因为监督工人的投入与产出是非常困难的。

几乎所有的劳工合同都允许雇主和雇员有一定的决策权。它们并不具体地规定在各种情况下各方需要做什么。雇主可以给工人分配不同的任务；[1]工人在一定范围内可以决定如何完成其任务。这种自由决定的行为本身就是信息不完全的结果（公司不可能预测到使利润最大化的任务应该分配给哪些工人）。即使合同的规定很具体，例如，规定了在特定的情况下分配什么样的任务，雇员也会发现很难确认某种情况是否发生了，也就很难确认公司是否遵循了合同的条款；类似地，雇员可以同意提供一定水平的努力程度，但公司会很难确认观察到的低产出是工人努力不够的结果，还是由于一些外生事件的影响。

本文关注的是雇主信息的缺乏和雇员可自由决定自己行为的情况。与之对称的关于雇员可自由决定自己行为的问题，该问题在某种意义上而言更简单，但某种意义上而言也更复杂。作为持久机构的公司需要建立声誉，因此雇员关心的是实际分配的任务的分布。我们假设在一个竞争性的劳动力市场

1　这点正是西蒙（Simon）关于雇佣关系的经典论文的核心要点。

上，公司吸引雇员的条件依赖于它已经建立起来的声誉。这是一个长期的分析视角，在短期只能通过雇员退出（以及法律的、联合的行动）的威胁来迫使公司遵守合同的隐含条款。然而，对市场这个方面问题的分析不在本文的分析范围之内。

1. 一般性模型

因为决定均衡工资合同的问题非常复杂，我们会考虑一些相关的、简化了的模型。但在开始分析以前，我们先给出一个一般性（但还不是最一般的）的模型，其后的分析就可以看作是这个模型的一些特例。

企业试图找到让利润最大化的合同，而该合同要满足工人接受它的约束条件。企业提供合同的各个条款不仅决定了谁会申请并接受该企业的工作，还决定了雇员在工作时候的表现。因此，我们以效用假设替换竞争性工资假设。[2] 在我们的假设中，企业都足够的小，使得它们认为自己面临的不同特征劳动的供给曲线是水平的，也就是说，它们提供的任一合同会使它们面临为零或者是无限多的劳动供给。而企业所面临问题的难点在于，无数可能的合同中的哪一个合同能够帮助企业实现最大化利润。

用更正式的方式来表达，我们假设企业里有大量的任务。我们可以把第 k 家企业将第 i 个任务分配给第 j 个雇员所得到的产出表示为：

$$ {}_kQ_i^j = {}_kF_i^j(e_i^j, \theta^j, \theta_i, \theta_i^j) \tag{1} $$

产出（每单位时间）是努力程度 e_i^j，——由雇员提供（每单位时间）——和其他一些变量的函数。这些变量可能和以下因素有关：（a）雇员的一些一般性的特征（如智商）θ^j；（b）一些技术性参数，θ_i（例如，任务的难度或面临的客观风险）；（c）描述雇员特征的参数，但该特征是和雇员在该任务上的表现相关联的，θ_i^j。

而雇员的努力程度又是该雇员和企业之间的合同（我们在此用 ω^j 表示）、雇员的信息结构（例如，关于他自身的能力以及分配给他的任务的技术参数信息。我们用 Ω^j 表示其信息结构）以及其效用函数的函数。

---

2 我们没有用更具有一般性的合同假设来代替工资假设的原因是：很明显地存在着无数可能的合同，在任一时间点只有一些合同在市场上是可得的，并且一般而言不存在可以用来对各种预期条款进行评估的隐性价格（implicit price）。因此，公司不仅要考虑现在被提供的合同中哪个是最好的，还要考虑在所有能被提供的合同中哪个是最好的。

这种分析均衡合同的框架最先是在分析佃农经济中被使用，斯蒂格利茨［参考文献 32］。在佃农经济中，合同不单单具体规定了工资，还包括了固定工资、分配比例、每个佃农耕种的土地的多少以及投入的劳动力的量——如果这些是可观察到的话。

为了简化分析，我们假设雇员的信息结构不受其行为（除了对工作的接受）的影响，但可能是时间的函数。给定该公司的工资合同和其他可得的工资合同，假设雇员需要解决的是一个最大化一生的期望效用的动态规划问题。这将决定其在工作时的每个时点上的努力程度：

$$e_i^j(t) = e_i^j(\omega^j, \Omega^j(t)) \tag{2}$$

企业关心的是最大化它的期望利润。[3] 期望回报 R 是各个任务产出的函数：

$$R_k = R_k({}_kQ_i) \tag{3}$$

此处[4]

$${}_kQ_i = \sum {}_kQ_i^j \tag{4}$$

此处是对所有被分配了该任务的雇员进行加总。

劳动力成本是各种被提供并被接受的合同的函数，也是被雇佣的工人质量的函数，还是工人努力程度的函数。企业可能会提供几种合同供工人选择。提供给工人的各种工资合约是当时观察到的工人特征的函数，也就是说，如果 $\omega_k^{l_j}$ 是第 k 家企业向第 j 位工人提供的第 l 种合同，那么有：

$$\omega_k^{l_j} \in W_k(\hat{\theta}^j) \tag{5}$$

其中$\hat{\theta}^j$为观察到的工人的特征。当且仅当信息结构为 Ω^j的工人在某家企业的工作上获得的一生期望效用大于从其他企业所能获得的效用时，该工人才会选择接受这家企业的工作。用 $W^j(\omega_k^l)$ 表示在企业 k 签订了合同 l 的工人所能获得的期望效用，那么：

$$W^j(\omega_k^{l_j}) \geqslant W^j(\omega_n^{m_j}) \tag{6}$$

对所有的 $\omega_n^{m_j} \in \omega_n(\hat{\theta}^j)$，所有的 m、n 成立。这决定了谁会向企业申请工作以及他们所申请的工作合同是什么样的。被观察到特征相同的工人可能有其他一些观察不到但对生产很重要的特征。有些这样的特征可以通过这些工人所申请的工作（合约）来进行推断。我们假设在某一合约下，企业雇佣到的工人是所有该工作的申请人的一个随机样本。因此，如果某一合约的申请人的平均质量要高于另一合约的申请人，那么我们假设该合约雇佣的工人的平均质量也要高于另一合约雇佣的工人的质量。

---

3 当公司为风险规避型的时候，很明显，需要做一些修改。

4 式（4）中包含的约束条件不是必需的（一般性的函数关系也可以）。

由此，我们可以把公司的期望利润简单地写为其提供的合约的函数：[5]

$$\pi = \pi(W_k(\hat{\theta})) \tag{7}$$

如果用 $W^*{}_k(\hat{\theta})$ 表示某企业提供的被一些工人接受的雇用合约集，那么，可以把企业的问题简单地表示为：

$$\max\pi(W_k(\hat{\theta})), \tag{8}$$

$$\text{其中 } W_k \in W^*{}_k(\hat{\theta})$$

而本文所要讨论的即是满足式（8）的解。

在我们大部分的分析中，我们进一步假设进入自由，这样在市场均衡时，最大利润为零。

2. 一般性模型的某些一般性推论：内部和外部劳动力市场的关系

本文的一个基本假设是：通过分析这里给出的竞争性市场均衡问题的解，可以找出哪些因素决定了被使用合约的性质。显然，要完全解出最大化问题所要求的信息和计算已经超出了许多市场参与者的能力范围，尽管如此，如果存在一个在很多方面明显优于其他的合约，那么"发现"了该合约的企业将会发现它们可以增加利润，而该合约也会被其他企业效仿。因此可以认为经济体具有逐渐移向我们所分析的合约安排的趋势。同时，考虑到"有限理性"，我们应该关注企业对相对简单的合约的选择。[6]

这个分析框架说明了有时被用在内部市场和外部市场间的二分法并不是非常准确。内部劳动力市场的特征在很大程度上是由外部市场决定的，我们可以从两个角度来理解这点：劳动力市场的结构影响了企业能吸引到工人的工资水平和工人的类型；其次，可行的合约性质依赖于以后其他雇佣机会，例如，通过分配给工人的任务的完成情况可以探知关于工人质量的信息，并且由于无法强制实施长期合约，那么工人在晚一些时候会获得同他们的能力相当的工资。

在接下来的几部分中，我们将从各个角度将这个模型具体化。在第二部分，我们只考虑取决于工人产出和投入时间（假设这是唯一能观察到的）

5　显然，有很多其他的重要变量，例如，资本存量也会影响到公司的利润。

6　然而，应该注意的是，对有限理性的讨论中所涉及的不完全信息问题不同于我们这里分析关注的问题。后者是指无法观察到某个参与人的某些特征，而前者是指关于特征（喜好、能力、机会）在总体中的分布信息的缺乏，以及准确地解决复杂的最大化问题的能力有限。

的劳动合约。我们先假设所有的工人都是同样的，[7] 唯一的随机变量是 θ_i——表示一种技术参数。公司只给每个工人分配一项任务。然后，我们考虑工人的能力（θ^j）不同以及能力越高的人越了解自己的能力的情况。在这种情况下，给定工资水平，哪些人接受该合约就变得很重要了（例如，式（6）所表示的关系就变得很重要）。第二部分的结论将考虑只存在单一能力，但需要分配不同任务的情况。

第三部分将分析监督者的引入是如何扩大了可提供合约种类的，例如，工资是除了时间和产出外某些变量的函数；并且根据工人能力的不同提供不同的工资，而不再是向每个工人提供同样的用工合约了。

二、计件工资和计时工资

许多用工合约的特点之一是工资和工作的时间以及工作的绩效都有关。我们称前者为计时工资。同工作绩效有关的工资或者采取在一定时间内支付更高报酬的形式（我们称之为计件工资），或者采取增加提升机会的形式（在未来某个时间获得更高工资的概率更高）。我们的目标是要找出是什么决定了计时工资和计件工资的相对重要性。此外，我们集中分析没有监督者的情况，在这种情况下，无法确知工人的努力程度，也无法直接观察到他的能力以及一开始假设未知的技术参数。因为我们考虑的是对计件工资和计时工资的选择问题，所以我们假设每个工人投入的劳动时间和其得到的物量产出（physical output）是可识别的。

按时间和绩效计算的回报一般都是非线性函数；例如，工人每周工作超过 40 小时，会收到更高的小时工资，而计件工资一般都规定了一个最低产量，使得绩效补偿函数如图 1b 所示。事实上，在纯定额制下，工人在完成了规定的工作量后获得固定的工资，如果产出增加也不会获得额外的报酬，因此纯定额制可以被看作是计件工资制的一种极端情况（如图 1c 所示）。在这里，我们只考虑线性计时工资制，[8] 并且要特别关注线性计件工资制，因为这两种制度分析起来都比较简单，而且在实际中也比较重要。

我们先找出影响对计件工资和计时工作的选择的几点因素：

1）如果主要的付酬方式是计件工资，那么工人收入的波动会增加，因

7 另外一种解释是每个工人是不同的，但他们的能力是同样的。

8 因为导致非线性计时工资制的主要原因同我们这篇论文所关注的问题非常不同。

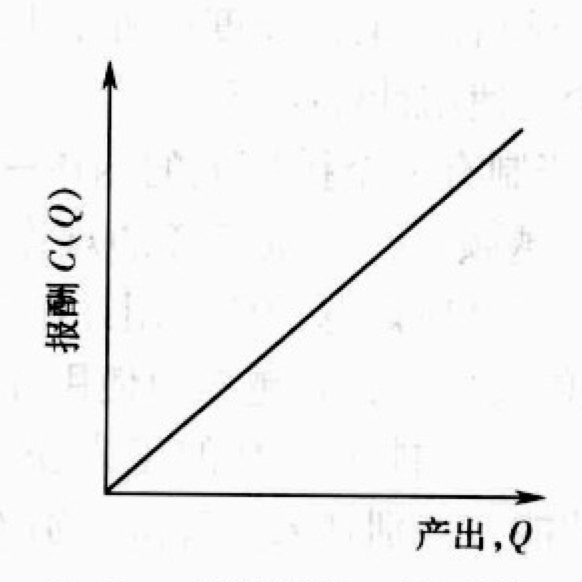


图 1a　线性计件工资系统

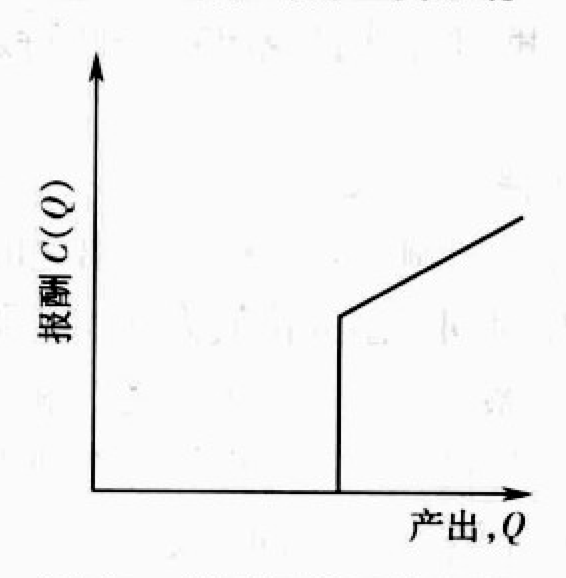


图 1b　有最低产量的报酬

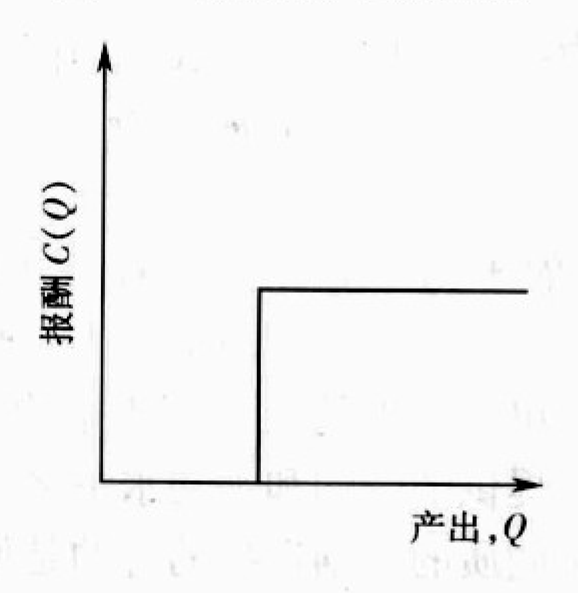


图 1c　纯粹定额制

此，如果风险越小，对风险的厌恶程度越低，则更多地使用计件工资。[9]

2）由于计件工资奖励的是速度，而没有对确保质量和爱护设备提供适

9　显然，降低工人的风险就会增加公司的风险。风险厌恶程度的不同是关键。然而，我们假设公司是风险中性的，所以合约由工人的风险厌恶程度决定。对此问题更一般的分析请参见斯蒂格利茨［参考文献 32］。

当的激励，[10]因此由工人控制的、工作速度以外的其他变量越重要，观察这些变量越困难，则要更少地使用计件工资。

3）由于使用计件工资制有一个操作上的困难——如何决定恰当的计件工资率，所以当技术变化越频繁，确认任务的难度（或者确认每个操作过程所增加的价值）越困难，就应该越少地使用计件工资制。

4）工人对货币激励的反应越大，越多地使用计件工资制。

5）由于在计件工资制下，能力越高的工人获得的报酬也越高，因此，当工人们完成某个任务的能力差别越大，直接识别该类能力越困难，则要更多地使用计件工资。

在接下来的几节中，我们尝试着找出在一些特殊情况下，均衡合约的具体决定因素。

1. 线性合约和风险厌恶型工人

θ 为风险变量，表示任一影响努力水平和产出之间关系的因素。e 为工人工作时的努力程度。假设单位时间产出的价值为 θ 和 e 乘积的一定比例，并且我们选择一定的时间单位使比例系数为 1（为了简化分析，假设没有使用其他要素，这样产出和边际生产率是一样的）。线性的激励机制支付给工人计时工资 I，再加上计件工资率为 r 的计件工资。工人的收入 C（每单位时间），为：

$$C = I + r\theta e \tag{9}$$

但在进入自由的竞争性均衡下（用上划线表示均值）：

$$\overline{C} = I + r\,\overline{\theta e} = \overline{\theta e} \tag{10}$$

因此：

$$C = r\theta e + (1 - r)\overline{\theta e} = r(\theta e - \overline{\theta e}) + \overline{\theta e} \tag{11}$$

单个工人的收入是给定合约下所有工人的平均边际产出和工人自己的边际产出的加权平均。或者说，单个工人的收入等于平均边际产出加上计件奖金，而该奖金等于该工人自己提供的劳动和平均水平之间的差值。

如果所有的工人都是同质的，则选择的合约是使工人期望效用最大化的合约。单个工人的期望效用是其收入和努力程度的函数，进而也是计时工资率和计件工资率的函数：

---

10 如果雇员需要自己支付自己所用的设备成本，则就不存在这个问题了。因此这个问题的严重程度取决于我们如何安排生产过程。在农业中也存在类似的现象：如果我们雇人锄杂草，按工人所锄杂草的重量来支付报酬，他们就会锄去最大的杂草而不是最有可能影响到蔬菜生长的杂草；如果我们雇工人种植蔬菜，按他们所种蔬菜的产出来付酬，他们就会恰当地锄草了。参见斯蒂格利茨［参考文献 32］。

$$W = EU(C, e^2) \tag{12}$$

此处，$U_1 > 0$，$U_{11} \leqslant 0$，也就是，假设工人是风险厌恶的，$U_2 > 0$，$U_{22} \geqslant 0$。在接受工作以后决定自己努力程度以前，如果该工人发现了 θ 的值，则当其的计件工资率为 r、计时工资率为 I 时，其会选择 e 满足下式：

$$U_1 r\theta - U_2 = 0 \tag{13}$$

如果在知道 θ 值以前就要决定努力水平，则有：[11]

$$E(U_1 r\theta - U_2) = 0 \tag{13a}$$

从式（13）或式（13a）解出 e，该式是 r（r 和 θ）的函数。将 e 代入式（12）。因此我们可以把期望效用写为计时工资率 I 和计件工资率 r 的函数：

$$W = EV(r\theta, I) \tag{14}$$

此处 V 是间接效用函数。[12] 均衡合约是在零利润约束式（11）下最大化式（14）的合约。经过一些处理后，如果 e 是在知道 θ 之前决定的，一阶条件得到的结论是：[13]

$$\frac{1-r}{r} = \frac{-\dfrac{E(U_C - \bar{U}_C)(\theta - \bar{\theta})}{\bar{\theta}\bar{U}_C}}{\eta} \approx \frac{R{s_\theta}^2 r}{\eta} \tag{15}$$

此处 s_θ 为 θ 的变异系数，R 为风险规避程度的阿罗—普拉特（Arrow-Pratt）度量，$-U_{CC}C/U_C$ 和 η 是努力补偿供给弹性 $(\partial \ln e/\partial \ln p)_{\bar{U}}$。另外，用 s_C

---

11 在本部分中，使用哪一个公式无关紧要，但对非线性的计件工资，则需要考虑其中的差别。

12 在式（13a）的情况下，V 是修正的间接效用函数，因为 e 将独立于 θ。

13 我们最大化：

$$EU(e(r,I)\theta r + I, e)$$

需满足以下约束条件：

$$I \leqslant (1-r)e\bar{\theta}$$

用 λ 表示约束条件的拉格朗日乘子，有：

$$eEU_C\theta - \lambda\left(e\bar{\theta} - (1-r)\bar{\theta}\frac{\partial e}{\partial r}\right) = 0$$

$$EU_C - \lambda\left(1 - (1-r)\bar{\theta}\frac{\partial e}{\partial I}\right) = 0$$

因为

$$dEU = EU_C dI + (EU_C\theta e)dr$$

$$\frac{\partial e}{\partial r} = \left(\frac{\partial e}{\partial r}\right)_{\bar{u}} + \frac{eEU_C\theta}{EU_C}\frac{\partial e}{\partial I}$$

代入，得到式（15）。为了得到近似值，我们发现：

$$\frac{EU_C(\theta - \bar{\theta})}{EU_C\bar{\theta}} \approx \frac{U_{CC}\bar{\theta}er}{U_C}\frac{E(\theta - \bar{\theta})^2}{\bar{\theta}^2} = \frac{U_{CC}\bar{C}}{U_C}\frac{re\bar{\theta}}{e\bar{\theta}}s_\theta^2$$

表示消费的变异系数

$$(1-r)=\frac{-\dfrac{E(U_C-\overline{U}_C)(C-\overline{C})}{\overline{C}\overline{U}_U}}{\eta}\approx\frac{Rs_C{}^2}{\eta}\qquad(16)$$

计件工资率只决定于的 θ（或者 C）与消费的边际效用的标准化协方差和努力供给弹性的比率。而标准化协方差[14]取决于相对风险厌恶系数的大小。所以从式（15）和式（16）可以得到以下结论：风险越低，风险厌恶程度越低，努力的供给弹性越大（激励效应越大），则计件工资率越高。我们注意到当工人是风险中性的，或者当 θ 没有变化时，$r=1$，工人获得的报酬就是其边际产出。如果相对风险系数在 1 左右，补偿供给弹性为 0.5，消费的变异系数为 0.4，那么大约有 1/3 的收入将来自于计时工资，而 2/3 来自于奖金。如果知道了 θ 的值后再决定努力程度，也可以得到类似的结论。[15]

14　当效用函数是二次函数时，R 为边际效用函数在收入均值时的弹性，则式（15）的近似值就是精确值。

15　替代式（15）的是：

$$(1-r)=-\frac{\dfrac{E(\beta-\overline{\beta})(C-\overline{C})}{E\beta EC}}{\overline{\eta}}\qquad(15a)$$

此处

$$\beta=\frac{V_1}{\lambda}+(1-r)\frac{\partial e\theta}{\partial I}$$

是调整了收入的增加导致的公司利润的下降后的收入边际效用，并且

$$E\beta=1\qquad\eta=\left(\frac{\partial\ln e}{\partial\ln r}\right)_{\overline{u}}$$

为补偿努力供给弹性，并且

$$\overline{\eta}=\frac{Ee\theta\eta}{Ee\theta}$$

为补偿供给弹性的加权平均。在一些简单的情况下，式（15a）可以获得显解。假设劳动的边际负效用为常数，收入的边际效用的弹性为常数，θ 的分布为对数正态分布，以及其变异系数为 s_θ^2。则：

$$(1-r)=\frac{1-(1+s_\theta^2)^{-\frac{1}{R}}}{\overline{\eta}}=\frac{s_\theta^2}{(1+s_\theta^2)\overline{\eta}}$$

当 $R=1$（效用函数为对数效用函数）时。如果 θ 的变异系数为0.4，并且 $\eta=1/2$，那么平均而言工人的收入有大约70%来自于奖金，30%是计时工资（比在知道 θ 之前就决定 e 的情况下要稍微少些）。

倘若努力程度对计时工资率增加的反应和 θ 的关联不是很高，我们可以将式（15a）近似地表达为式（16）或者

$$\frac{1-r}{r^2}\approx\frac{Rs_{Y2}}{\overline{\eta}}$$

此处 s_Y 为生产率（$e\theta$）的变异系数。

2. 计件工资率和自选择

当不同能力的工人愿意接受某个合同时，至少有两个理由使得公司倾向于只雇佣最有能力的工人：只要工人接受的工资有一部分是计时工资，也就是他们的收入有一部分是同绩效无关，那么平均的劳动力成本将取决于工人的平均生产率。即使所有的工人获得的报酬都是他们的边际产出，如果有特殊的培训（雇佣/解雇）成本，那么在发生这些成本之前知道被雇工人的质量就很重要了。最近出现了大量关于甄别——根据能力将个体分类的过程——理论的文献。[16]这些文献大多数都关注于如何使用考核和如何使用能力的替代变量，例如，受教育程度。本节的目的是证明计件工资制在甄别中的作用，以及证明能力较低的工人的存在会导致更高的计件工资率。

考虑最简单的可能情况：只存在两组不同能力的工人，工人都知道自己的能力。在图 2a 中我们画出了他们的保留工资线（reservation-wage curve），该曲线为计时工资率和计件工资率的函数。如我们所画的曲线所示，如果所有的报酬都是以计时工资的形式支付的（这意味着更胜任该工作的工人的机会成本更大，也就是，完成该工作的能力和完成其他工作的能力是相关的），能力较低的工人的保留工资也较低，但如果所有的报酬都以计件工资的形式支付的（意味着在完成这个特定任务时能力低的工人有相对的劣势），保留工资会更高。图 2b 和图 2c 描述了其他一些可能的情况，图 2b 描述了当所有的工人都有相同的时间机会成本时的情况，图 2c 描述了能力较低的工人在完成给定任务时仍可能有相对优势的情况。

企业雇佣了一位类型为 j 的工人的期望利润为[17]

$$Ee^{j}\theta^{j}-rEe^{j}\theta^{j}-I$$

图 3 画出了等利润线。如果经济体中只有生产率较高的工人，则均衡点为点 M_1，这在前一节中已经描述过了，我们可以在图 2a 中标出相应的点。当只有生产率较低的工人时也可以得到类似的分析。图 2a 中低能力工人的保留工资曲线上的 M_2 点最大化了期望利润。

现在，企业面临提供哪个合同的选择。它可以尝试提供一个能同时吸引

16 例如参见 Stiglitz［参考文献 28，33］，Rothschild and Stiglitz［参考文献 20］，Spence［参考文献 23，24］，Arrow［参考文献 4］，Salop and Salop［参考文献 21］，Akerlof［参考文献 1］。

17 可以把 θ^{j} 看作是一个综合随机变量，表示所有其他随机变量对第 j 个工人每单位努力的产出的影响。

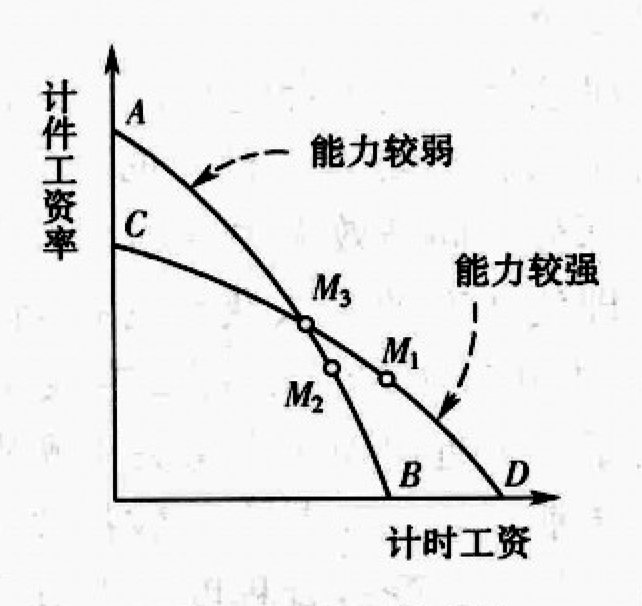


图 2a　保留工资曲线

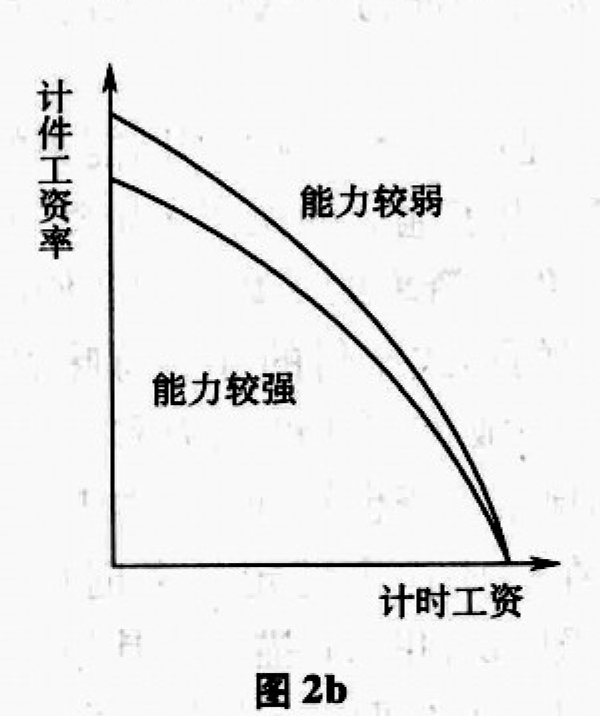


图 2b

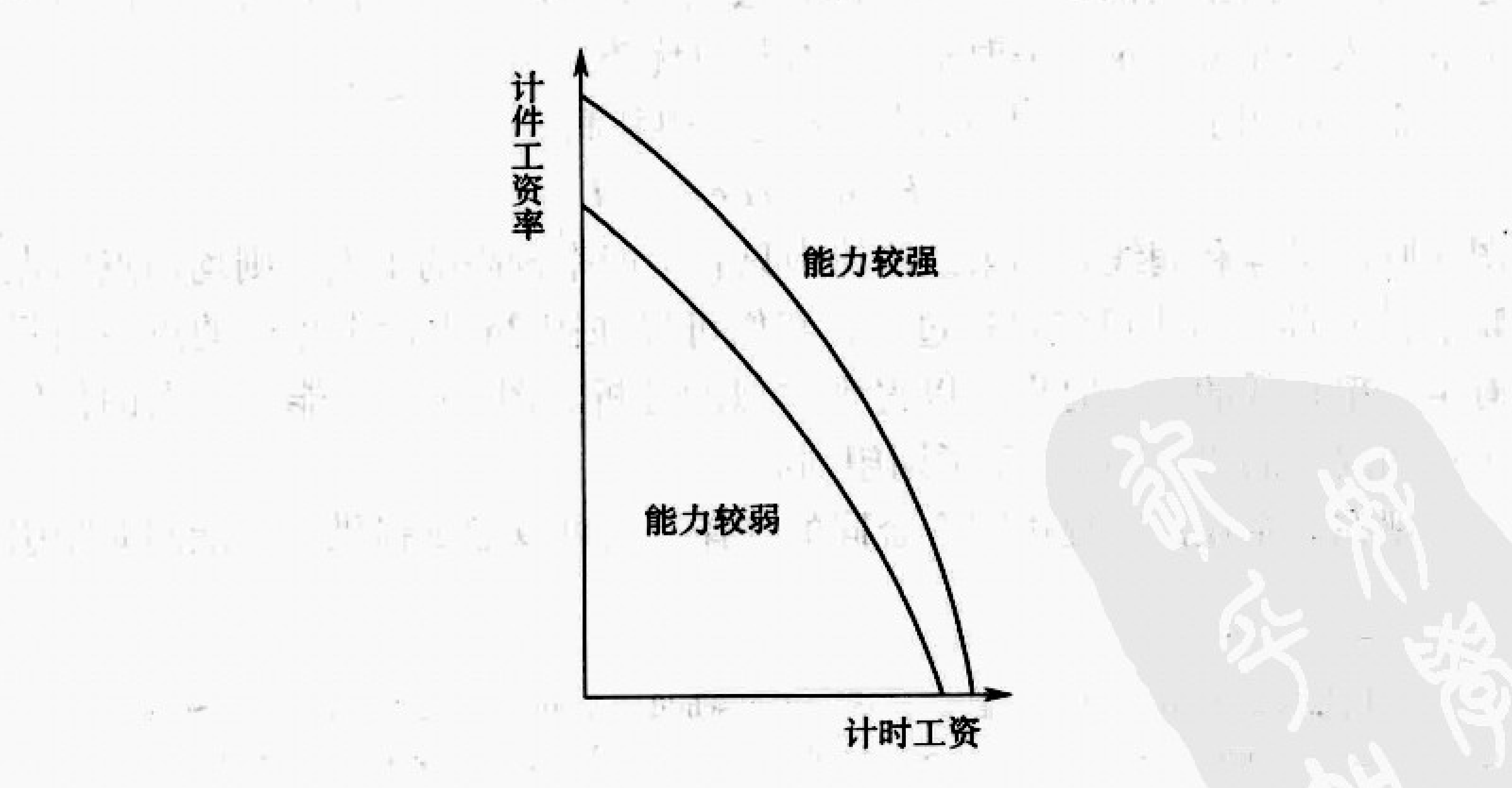


图 2c

两组工人的合同，或者提供一个只吸引一组工人的合同。例如，考虑图 2a 的情况，M_3是两条保留工资线的交点。曲线 M_3C 上的任一合同都只会被能力较高的工人所接受，而在曲线 BM_3 上的合同只会被能力较低的工人所接受。在曲线 AM_3D 上的任一合同都会被两组工人所接受。如果公司只希望雇用一种质量的工人，则最优合同就非常明显了：它或者提供 M_2，只得到低质量的工人，或者提供 M_3（CM_3 上的、只比 M_3 的计件工资率稍微高一点的合同），只得到高质量的工人。因此，要想雇用质量较高的工人就得提供较高的计件工资率。

但是也有可能公司发现它自己处于这样一种处境：它不能（如图 2c）或者不愿将能力较低的工人排除掉。那么能力高的工人的保留工资线上任一点的利润率都是两组工人利润率的加权平均，权数为两组工人的相对规模（在所有申请人中）。[18]通过增加计件工资率，虽然能力较高的工人带来的利润率会降低，但能力较低的工人带来的利润会上升（假设供给弹性的差别不大）：再次推测会有更高的计件工资率。这点可以从图 3b 中看出。因为能力较低的工人的等利润线比能力较高的工人的等利润线更陡峭，[19]对 M_1 点的计件工资率的一点点增加会对来自于高能力工人的利润产生轻微影响，但对来自于低能力工人的利润有很大的正向影响。我们注意到因为能力较低的工人都在他们的保留工资以上，因此只有高能力工人的保留工资需要考虑。

更一般地，对公司宣布的每种计件工资率和计时工资率，都有不同的申请者。公司的利润可以被写为：

18　Rothschild and Stiglitz［参考文献 20］的分析证明了：一般而言，不存在高能力工人和低能力工人一起工作的均衡，也就是，如果我们把他们对保险市场的分析用到我们这个问题上，则唯一的均衡为 M_3，即商品的价格会调整到使合同的利润为零。然而，可以认为这里的分析是在更一般的条件下进行的，例如，不一定是自由进入的情形，所以利润可能不为零，并且没有两个公司是完全一样的。因此，每个公司可以把保留工资曲线视作是给定的。

19　把等利润曲线定义为：

$$\bar{\pi} \equiv (1-r)Q - I$$

此处的 Q 为其平均产出。因此，

$$-\left(\frac{\partial r}{\partial I}\right)_{\bar{\pi}} = \frac{\dfrac{1}{Q} - \dfrac{(1-r)}{I}\dfrac{\partial \ln Q}{\partial \ln I}}{\left(1 + \dfrac{\partial \ln Q}{\partial \ln r}\dfrac{(1-r)}{r}\right)}$$

所以如果 $\partial \ln Q/\partial \ln I$ 和 $-\partial \ln Q/\partial \ln r$ 的值相差不大，则 Q 的值越小——生产率越低——等利润曲线越陡峭。直观地讲，计时工资率的增加使每个工人的成本增加都是一样的，不管能力高低，但生产率较低的工人的每单位产出的成本增加更多。因此，为了保持利润不变所要求的计件工资率的降低就更大。

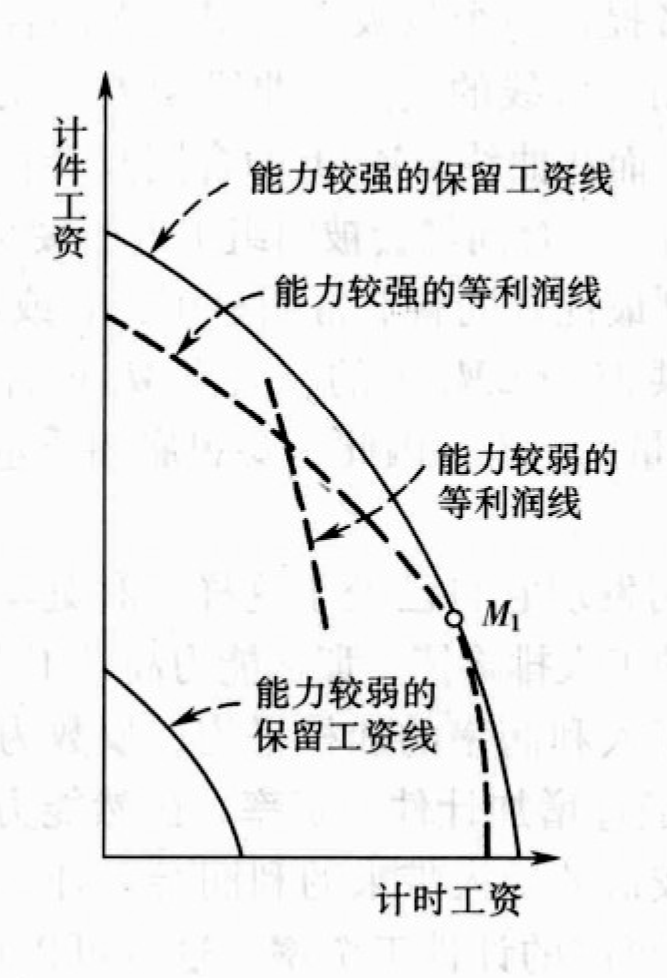


图 3 自选择均衡

$$Q(1-r)-I$$

此处 Q 为每个工人的平均产出，也就是 $Ee\theta$，这里不仅对每种申请者的 θ 求期望，也对申请者类型求期望。因此，内点解将满足：[20]

$$(1-r)\frac{\partial Q}{\partial r}-Q=0$$

$$(1-r)\frac{\partial Q}{\partial I}-1=0$$

而产出均值 Q 的变化不仅和激励效应（r）、收入效应（I）有关，也同自选择效应（self-selection effect）——申请者构成的变化——有关。一般而言，这样将会增强计件工资率的增加对平均产出（每个工人的）的影响；而现在 I 的增加可能会增加 Q（而不是降低）；也就是说，如果能力更高的工人的时间机会成本更高，则 I 的增加会提高申请工人的平均质量。这点确保了内点解的存在，在内点解处，I 的进一步减少（r 给定）不会使申请者的数

---

20 这里的分析再次使用了局部均衡的方法。如果存在大量相同的公司，则均衡又成为一个难题，但不仅仅是因为脚注 18 中提到的那些原因。即使工资成为公司可以用来影响其工人质量的唯一工具，也可能不存在均衡（Rothschild-Stiglitz 的分析假设有两种或更多的工具可以使用，如价格和覆盖率），参见 Stiglitz［参考文献 31］。

目变为零。这对应着一个有名的现象——公司会支付高于它们需要支付的报酬来吸引更高质量的工人。而这意味着均衡不一定使得供给等于需求：传统的认为当劳动力的供给大于需求时，工资会降低，使市场均衡的观点在这里不再适用。[21]

然而，我们注意到：如果工人完全知道他们自己的能力，并且没有其他的风险，那么均衡将保证公司能够设定 $r=1$，并将工作机会拍卖给出价最高的竞拍者，也就是，出现非正的计时工资（如果计时工资是正的，则公司会蒙受损失）。这个结论独立于总体中工人的能力或者对工作的偏好的分布，出价最高的工人不一定是能力最高的，出价取决于工人相对于自己的机会成本的执行任务的能力，如果所有的工人都有一样的机会成本，那么就是最有能力的人获得工作。[22]

3. 多种任务和线性激励制

前两节讨论的工作都可以被视作是“单任务工作”：工人接受一个工作，该工作使工人执行一个任务，而该工人不知道在执行这个任务时需要投入多少努力。

然而，许多工作都需要工人完成多种任务。例如，机械工厂里的工人需要制造各种各样的机械。计件工资率可能不同，但并不同工人完成每个任务所需要投入的努力成比例。如果有大量的任务需要完成，工人可以通过借贷活动使得从执行每个任务中获得的收入的边际效用近似相等（也就是说，他可以平滑收入流），但是所需要的努力不同。因此在决定均衡工资合约时起关键作用的是“努力风险规避程度”，而不是“收入风险规避程度”。而且，如果任务的分布是给定的，通过改变执行不同任务的速度，可以改变执行难度不同的任务所需花费的时间的分布。这将导致负的计时工资率的出现（当然负的计时工资率常常无法直接观察到，但通过设定一个低的“基本工资”和一个高的计件工资率，如图 4 所示的那样，将计件工资制反推到“起始点”，将出现负的计时工资率）。

这里给出的模型对第一部分的模型做了一点修改。我们假设所有的工人都是一样的，θ_i 表示一单位时间内在 i 任务上花费的一单位努力的产出，

21　Akerlof［参考文献 1］和 Stiglitz［参考文献 33］。

22　莫里斯（Mirrlees）的论文模型是一种特殊的情况：工人的能力不同，但时间的机会成本一样。他的结论——所有的工人获得的报酬都等于其边际产出——等价于最后只有一组工人被雇用的结论，而这可以用这里讨论的线性激励安排来达到。

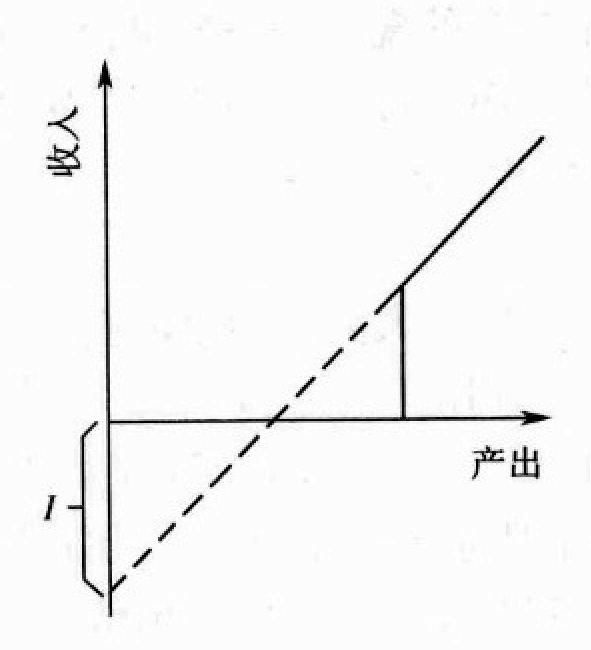


图4 隐含的负的计时工资率

因此$1/\theta_i$为当工人的努力水平为 e 时，完成该任务所需要花费的时间。工人在大公司工作，将分配给他的任务的分布视作是给定的，同他自己的努力水平无关。如我们前面所讨论的，在此采用以下的效用函数形式是合理的：

$$C - Z(e)$$

此处 Z（e）为努力带来的负效用。所以，当被分配了第 i 个任务，计时工资率为 I，计件工资率为 r（我们将单位标准化，使得所有任务的计件工资率相同）时，工人最大化：

$$I + re\theta - V(e) \tag{17}$$

此处我们去掉了 θ 的下标。为表达的方便，我们引入下一变量：

$$x = e\theta \tag{18}$$

表示生产的速度。对于单个工人取下式最大值：

$$I + rx - V\left(\frac{x}{\theta}\right) \tag{19}$$

也就是，[23]

$$r = \frac{V'}{\theta} \tag{20}$$

公司选择 I 和 r 来最小化其期望成本，因为生产一单位产品花费θ/e 时间，所以期望成本是：

23 这里假设对每个任务，工人在决定 e 值之前是知道 θ 的（或者他很快就知道 θ 了）。

$$IE\frac{1}{x}+r \tag{21}$$

在满足约束条件的情况下，取式（21）最小值，而约束条件是工人每单位时间的期望效用等于其机会成本 $\overline{W}$：

$$I+rEx-EV(x\theta)=\overline{W} \tag{22}$$

将式（22）代入式（21），再求微分（利用式（20）），我们得到：

$$1-ExE\frac{1}{x}-IE\frac{1}{x^2}\frac{dx}{dr}=0$$

或者

$$I\equiv-\frac{ExE\frac{1}{x}-1}{E\frac{1}{x^2}\frac{dx}{dr}} \tag{23}$$

因为$1/x$ 为凸函数

$$E\frac{1}{x}<\frac{1}{Ex} \tag{24}$$

所以式（23）的分子为正，从式（20）得：

$$\frac{dx}{dr}=\frac{\theta^2}{V''}>0 \tag{25}$$

所以，最优的激励机制使得 $I<0$。为了得到式（23）中一些项的数量级，我们做泰勒级数近似，假设 $V(e)=ke^{\alpha}$，以及 θ 为对数正态分布，再代入变异系数 s_θ，则有下式：

$$\frac{I}{Erx}=\frac{-\alpha[(s_\theta^2+1)^{(\frac{1+\alpha}{\alpha})^2}-1]}{(s_\theta^2+1)^{(\frac{1+\alpha}{\alpha})^2}}$$

θ 的方差越大，努力的边际负效用的弹性越大，则 I 的负值程度越大（相对于总报酬，倘若 $\alpha\geqslant1$）。

对计件工资制的修改

通常工人们对计件工资制的一个抱怨是，如果他们工作更努力，计件工资率就会被修改。因此我们的问题是如何来解释这种修改程序。当然，如果能够很容易地判断工人产出的价值，就会简单地根据其边际产出支付其报酬。但在以下两种情况下，是无法确认工人产出的价值的：

1）工人以小组的形式工作，而只能确认小组的产出。[24]

2）可以确认工人的物质产出，但没有价格可以帮助公司决定其产出的价值。在专业化分工生产中这是比较典型的情况：最终产品有市场价格，但生产过程中每个阶段或环节的产出没有价格。

在这些情况下，工人可能比经理知道更多关于生产的真实成本（例如，工人付出的努力的量）的信息。这也可以被视作是一个甄别问题：公司试图从工人那里获得关于劳动力需求的真实信息，用以根据工作的难度，把不同的工作分类。当工作很难时（计件工资率太低时），工人有很强的动力向经理提供信息，但当工作比较简单，也就是计件工资率较高时，工人没有相应的动力向经理提供信息。因此，直接来自工人的关于分配工作难度的信息缺乏可信度。公司试图通过工人的行为来判断任务的难度，当它获得了该信息时，就会用该信息来修改原来的计件工资率。

当信息缺乏时，不同难度的工作的计件工资率是一样的，这就会存在扭曲：消费（收入）和努力的边际替代率在不同的任务上是不同的。虽然信息可以降低这种扭曲，但这种通过工人过去的表现来推断任务难度的方法是有成本的，它降低了对工作的激励，实际上，它降低了计件工资率。我们再次得到这样的结论：在设计最优合同时，要考虑到获得信息的边际收益和边际成本。

均衡合同是在外生给定工人的期望效用水平上最大化期望利润的合同。我们想要证明的是：在简单的线性修改制下，对计件工资率的一些修改通常对公司来说是必要的。

因此，我们考虑前一节所用的模型，但这里假设任务是"成批"分配的，例如，要求工人生产 10 个某种类型的机器。公司可以宣布在机器的一部分完成以后，或者在给定的一段时间过去以后，它会根据工人之前的表现，按照一定的规则调整计件工资率。

我们需要区分两种不同的情况：一种情况是一个任务只分配给一位工人。由于无法比较该工人和其他工人的表现，公司无法确定该工作的真实难度。则对计件工资率的修改只能参考该工人自己的产出。另一种情况是有一大组的工人在完成相同的任务。修改则参考的是该组的平均产出，同时单个工人的收入将取决于他自己的产出所带来的计件工资。我们假设该组足够大，使得单个工人忽略了自己的产出对修改的影响，从分析的角度看，第二

24　这方面的内容在 Rothschild-Stiglitz［参考文献 20］和 Akerlof［参考文献 1］中有所讨论。

种情况更容易分析一些，所以我们先来讨论这种情况。

线性修改制和大组

对这种情况的分析比较简单，因为从单个工人的角度看，两期（修改前和修改后）是完全分开的。

我们假设在第一期工人的收入是

$$I_1 + r_1 x_1$$

此处和前面提到的一样，x_1 为生产的速度，也就是

$$x_1 = e_1 \theta$$

在第二期，工人的收入为

$$I_2 + r_2 x_2$$

此处假设 r_2 是第一期表现（x_1）的线性函数

$$r_2 = \beta - \gamma x_1 \tag{26}$$

因此，在第 i 期，工人让下式最大化：

$$I_i + r_i x_i - V\left(\frac{x_i}{\theta}\right) \tag{27}$$

有，

$$\theta r_i = V'\left(\frac{x_i}{\theta}\right) \tag{28}$$

公司则要最小化其期望成本 C。用 N 表示每个被执行任务所包含的工作量。那么在修改发生后，还有$(N - x_1)$的工作量需要完成。如果工人以 x_2 的速度工作，将花费该工人

$$l = \frac{N - x_1}{x_2} \tag{29}$$

的时间来完成该任务。因此，期望成本可以被写为：

$$EC = E[I_1 + r_1 x_1 + I_2 l + r_2 (N - x_1)] \tag{30}$$

公司最小化式（29），满足以下的约束条件：支付的报酬要足以使其能够雇到工人，也就是，工人的期望效用（每单位时间）必须等于机会成本

$$E\frac{C - V\left(\frac{x_1}{\theta}\right) - V\left(\frac{x_2}{\theta}\right) l}{1 + l} = \overline{W} \tag{31}$$

具体的计算[25]证明了，如果我们使第一期的时间足够长，则

$$\frac{dEC}{d\gamma_{\gamma=0}} < 0 \tag{32}$$

也就是，对计件工资率进行一些修改总是合算的。[26]

修正制的一个有趣的推论是：在第一部分分析的情况下，工人联合起来更努力地工作对他们是有好处的，但在这里却恰恰相反，工人们相信，如果他们串通起来不努力工作对他们是有好处的。如果他们把薪酬参数看作是给定的，他们会观察到通过串谋，可以增加自己的期望效用。他们相信工作努力的工人会给其他工人带来负的外部性。但如果他们真的串谋，而且如果劳动的供给是有弹性的，针对串谋行为，公司会改变薪酬制的参数。公司将观察到努力的供给函数不是由式（28）的解给出，而是由下面的式（34）和式（35）的解给出，也就是，工人考虑了努力水平对修改行为的影响的情况。在新的均衡下，串谋的结果是，工人的期望效用还是保持原来的水平不变，但期望的生产成本增加了。一旦公司调整薪酬制会将串谋考虑在内，所

---

25 没有修改时，$\gamma=0$，$I_1=I_2$，$r_1=\beta$。因此

$$\left(\frac{dEC}{d\gamma}\right)_{\gamma=0} = \frac{dI_1}{d\gamma} - E\frac{I_2}{x_2^2}(N-x_1)\frac{dx_2}{d\gamma} - Ex_1(N-x_1)$$

从式（28）直接得到 $dx_2/d\gamma<0$。我们前面的分析证明了当 $\gamma=0$ 时，$I_i\leqslant 0$。最后，将式（31）重写为

$$\Phi \equiv E\frac{I_1+I_2l}{1+l} + \frac{(r_1x_1 - V(x_1/\theta))}{1+l} + \frac{l[\beta x_2 - \gamma x_1x_2 - V(x_2/\theta)]}{1+l}$$

当 $\gamma=0$

$$\Phi_\gamma = E - \frac{x_1x_2l}{1+l} = -E\frac{x_1(N-x_1)}{1+l}$$

$$\Phi_{I_1} = E\frac{1}{1+l}$$

因此

$$\left(\frac{\partial I_1}{\partial\gamma}\right) = \frac{E\dfrac{x_1(N-x_1)}{1+l}}{E\dfrac{1}{1+l}} < Ex_1(N-x_1)$$

如果第一期足够长，使得

$$\frac{dx_1(N-x_1)}{d\theta} = (N-2x_1)\frac{dx_1}{d\theta} < 0$$

因为 $d[1/(1+l)]/d\theta>0$，也就是，如果初始期足够长，使得至少一半的任务是在修改之前完成的。

在上面的计算中，求期望的 θ 的值足够小使得任务不会在第一期就被全部完成了。

26 即我们可以找到第一期的一定长度使得对一些 θ 的值，做修正是必需的。

有的工人会观察到：如果他们串谋不成功，则期望效用将低于他们的机会成本。[27]

线性修改制和单人工作

这里我们考虑分配给每个工人的工作都有所不同的情况。在这种情况下，工人清楚地知道他们在第一期的努力程度对他们第二期的报酬的影响，和对第二期所用时间长度的影响。

因此，工人选择 x_1 和 l 以最大化：

$$W = \frac{I_1 + I_2 l + r_1 x_1 + \beta(N - x_1) - \gamma x_1 (N - x_1) - V\left(\frac{x_1}{\theta}\right) - V\left(\frac{x_2}{\theta}\right) l}{1 + l} \tag{33}$$

也就是

$$r_1 - \beta - \gamma(N - 2x_1) - \frac{V'\left(\frac{x_1}{\theta}\right)}{\theta} + \frac{V'\left(\frac{x_2}{\theta}\right)}{\theta} = 0 \tag{34}$$

和

$$I_2 - W + \frac{V'(N - x_1)}{\theta l} - V = I_2 - V\left(\frac{x_2}{\theta}\right) + V'\left(\frac{x_2}{\theta}\right)\frac{x_2}{\theta} - W = 0 \tag{35}$$

公司再一次希望在服从于式（31）的约束条件下，让期望成本最小化，但这里的 x_1 和 l（或者 x_2）由式（34）和式（35）给出，而不是式（28）。再次可以证明得到：

$$\frac{\partial EC}{\partial \gamma_{\gamma = 0}} < 0$$

27　我们假设在工厂里有默认的或公开的串谋，但仍存在一个竞争的劳动力市场，在职的工人不能限制对公司的劳动力的供给。

如果第一期的时间足够长。[28]

三、关于等级理论

最近对等级生产结构（hierarchical production structure）现象的一些解释暗含的理论基础是经济学的“阴谋论”（conspiratorial view），即不知何故，公司老板建立了使老板的存在成为必要的生产结构。虽然这其中有很多值得商榷的问题，但我们在这里不想评价该种观点正确与否，或者质问是否能够建立一个与该结论相一致的“内生技术变化模型”。而我希望提出另外一个模型来解释，在这个模型中，等级生产关系是竞争力量自然而然导致的结果。

我们需要区分“等级”这个词的两种不同意思：（a）垂直等级，也就是，使用权力（直接的命令）而不是货币的或其他形式的激励，（b）水平等级，也就是，对工人使用不同的待遇，如地位、报酬，等等。这里，我们只分析前者。

垂直等级一定会导致监督者的存在。监督者被视作获得了原来缺乏的信息，监督者的存在是否有经济效率，取决于该信息相对于监督者成本的价值有多大。

监督者的另外一些功能和其决策制定者的角色有关。我们也讨论了在信息的一些非凸性的基础上，应该使决策权集中化。

1. 监督者和不完全信息

我们在第二部分的开头列出了影响计件工资和计时工资的选择的主要因

---

28 与之前的计算相比唯一的变动来自式（35），我们可以解出 x_2/θ，独立于 γ。因此，$\partial x_2/\partial\gamma=0$。在 $\gamma=0$，$r_1=\beta$ 时

$$\frac{\partial x_1}{\partial\gamma}=-\frac{(N-2x_1)}{\frac{V''}{\theta^2}}$$

$$\frac{\partial x_1}{\partial\theta}>0,\frac{\partial x_2}{\partial\theta}>0,\frac{\partial l}{\partial\theta}<0$$

$$\frac{dEC}{d\gamma}=E\left[I_2\frac{\partial l}{\partial\gamma}-x_1(N-x_1)+\frac{\partial I_1}{\partial\gamma}\right]$$

$$\frac{\partial I_1}{\partial\gamma}=\frac{E\frac{x_1(N-x_1)}{1+l}}{E\frac{1}{1+l}}$$

其余的计算同前面一致。

素。使用计件工资制的理由中有两条是和信息的缺乏有关的，而监督者可以在这方面做部分的补救：（a）由于无法直接观察到投入而需要激励（这种情况下，监督者可以监控投入）；（b）当工人的质量无法直接观察到时，用计件工资制来影响劳动力的质量（这种情况下，监督者可以监控工人的能力）。同时，也可以用监督者来减轻计件工资制的两个局限性：（a）监督者可以监控产出，例如，确保产出达到质量标准，或者当只有小组产出能被直接观察到时，[29]确认各小组成员的贡献；（b）监督者需要确定合适的计件工资率。因此监督者是替代了价格体系（价格体系可以提供合适的计件工资率）的作用。

下面，我们将详细说明监督者作为投入监控者的作用。第二部分证明了如果无法直接观察到工人的努力程度，则均衡合同将由计时工资和计件工资共同构成。产出是随机的，而工人的投入又是不可识别的，因此工人无法获得对不确定收入的完全保险。同时，他获得的部分保险又会降低工作的积极性。因此信息的不完全导致了两种成本：工人承担了更多的风险，而其工作的努力程度也更低。关于投入的信息也可能是可获得的，但需要付出成本：支付监控工人投入的监督者的成本。

工人的工作有几方面是需要监督人监控的。当产出和 θ 可以被观察到时，可以推断出 e，而如果产出和 e 可以被观察时，可以推断 θ。这使得我们可以考虑另外一些合同。

最简单的合同是使工人同意投入给定水平的努力。此时产出是取决于 θ 和 e 的，但工人的报酬不取决于产出。如果这种合同是可行的，对工人而言（事前的概念），这将比他签订先前讨论的合同要好。因为这样他可以完全消除风险。最优合同则是：

$$\theta U'(C) = V'(e) \tag{42}$$

此处

$$C = e\bar{\theta}$$

如果 $U''' > 0$，$U'\theta$ 是 θ 的凹函数，

$$EU'\theta < U'(e\bar{\theta})\bar{\theta} \tag{43}$$

该合同下工人的平均收入要高于我们第一部分中给出的合同下的平均

29 Akerlof [参考文献 1] 和 Stiglitz [参考文献 33] 在讨论识别个人能力中的自选择机制时，已经充分地讨论了这一类的生产活动。

收入。[30]

因此，工人自愿要求被监督，而竞争均衡的一个特点是有一定量的强制。与在奖金制下提供的激励相比，他们工作得更努力。他们也愿意被迫更努力地工作，因为这样会有更高的期望效用。[31]

虽然每个工人都可能非常讨厌这种强制性，认为对自己是没有必要的，但他又偏好于在有这种强制制度的公司里工作，因为他意识到如果没有这样的强制制度，一些同事会怠工，此外使用一定程度的强制手段的公司还能够支付更高的工资。

当被监控的 θ 为外生决定的风险变量时（天气、工人工作时使用的机器的质量等），这种合同是合理的；但当风险变量是工人完成某个任务的能力时，能力更高的工人会倾向于假装自己并不是特别有能力，从而可以节省自己的努力。这就还需要有一定的激励机制。

在考虑是否接受带有这种强制条件的合同时，不喜欢权力关系的工人或许会要求较多的额外补偿，以便接受这种雇佣关系。这种关系的一个缺点是它不允许一种在计件工资制下存在的个体多变性（例如，每天的工作量不同）。这种合同的存在反映了，对一些工人而言，使用这种合同的公司所能提供的更高工资和确定收入能够补偿这点不足之处。

对于是否应该将本节定义的这种关系视为权力关系，还有一些含糊之处。像第二部分中讨论过的那些激励机制肯定不是权力关系。另一方面，极端非线性的激励机制——如果工人没有完成该任务，则得不到报酬或者会被解雇，可能被认为也可能不被认为是一种权力关系。如果对于工作是否能够被完成不存在不确定性，那么监督者没有必要做出决策，我们也不可能认为

---

30 为了保证 $U'\theta$ 的凹性，如果 e 是在 θ 之后决定的，则我们要求 $d^2e/d\theta^2>0$。

31 从事前的角度，一个更好的合同是工人承诺其投入的努力是 θ 的函数。如果那样的合同是可实施的，最优合同会使工人得到的报酬不变，不管其面临 θ 是多少，但经历更高 θ（例如被分配了更好的一块土地或者更轻松的工作）的工人会更努力地工作。即工人签订的合同将最大化：

$$EU(C(\theta))-V(e(\theta)) \tag{i}$$

此处

$$EC(\theta)=Ee(\theta)\theta \tag{ii}$$

式（i）的解使得 $U'=$常数，对所有的 θ 都成立，并且 $\frac{V'}{\theta}=$常数也对所有的 θ 都成立。θ 越高的工人努力的边际负效用也越高，也就是，他们工作得更多。他们事后的效用实际上要低于那些能力更低的工人。这些结论类似于莫里斯关于在完全管制经济体中，在功利主义框架下的收入的最优分配的结论。当 θ 表示工人完成某个工作的能力时，该合同带来的负面激励效应（也就是，工人会隐藏关于自己的能力的信息）比先前讨论过的合同的更大。

这其中存在权力关系。但如果任务是否能被完成存在不确定性，则监督者事实上要决定是否给工人报酬。结果，由于报酬将是随机的，监督的程度（或质量）将同时影响被使用的最优激励机制和工人将会得到的期望效用的水平。因此这里给出的分析和第二部分的分析都可以被看作是两个极端的例子——对投入没有监督的情况和对投入完全监督的情况。大多数的雇佣合同都包含部分监督，而不是完全的监督。[32]

2. 监督者和干预

监督者在决策，干预下属行为，下达命令等方面的作用或许被认为是对监督者作用的诠释的最核心部分。前一部分将监督者这样的角色描绘为一个合同的强制实施者。

然而，监督者作为下达命令者，作用要多于仅仅作为一个合同实施者的作用。经济体中的所有代理人在执行任务时都有一定的自主权，但另一方面，在有些情况下，监督者会对这种自主权有一定的干预。我在这里讨论的正是监督者这方面的作用。

我们可以想像，当新的或者不寻常的意外事件发生，或者面对新的信息时，公司可以采取多种应对措施：

（1）公司设计一套激励机制，使得当意外事件或新的信息发生时，工人会采取对公司最有利的应对措施，并且对工人没有直接的监管。

（2）公司根据各种意外情况的不同，设计出一套常规规则（命令），并设计出监督—奖金机制来促使工人遵循这些常规规则。

（3）公司指示工人当某些意外情况或某类新的信息出现时，要向监督者咨询，由监督者做出决策，并指导工人采取措施。

（4）公司对各种意外事件（或新的信息）进行监督，当需要工人采取应对措施的意外事件发生，而不是日常的激励机制（或日常的指令）可以预防的事件发生时，监督者就会进行干预并制定新的一套指令或者新的激励机制。

（5）最后，公司可以没有意识到这种意外事件，而将这种意外事件视作“自然状态”，就像它是日常自然状态中的一种情况一样。

前两种措施把监督者的作用局限在我们前面讨论过的方面；[33]第三种使得监督者的作用像一位咨询师；在第四种措施中，监督者直接干预工人的行

32 对这些问题的进一步的讨论，请参见 Stiglitz［参考文献 32］和 Mirrlees［参考文献 16］。

33 这种常规化很像马奇和西蒙（March and Simon）所称“行为规范”（performance programs）的概念。

为，正是这种作用才是权力关系最清楚的表现。因此，在第四种组织形式中，监督者会干涉工人在执行任务时的自主权。以上对特定意外事件采取的每种应对措施都既有成本也有收益。

我想在本节讨论的是这些信息流都有许多不同的非凸性，而这些非凸性导致了专业化，并且给权力关系提供了经济上的正当理由（当然，对组织形式的选择取决于许多其他因素，例如，工人对权威的态度。这些方面的问题在组织学的文献中有深入的讨论。[34]本部分的讨论可以被看作是对这类文献的补充，关注于在这之前没有被充分关注过的一些方面）。

收益

如果没有干预行为，则被监督者可能会采取一些不利于公司的行为，这就是干预行为的回报。当把行为都规则化的情况下，对反复发生的一些事情只存在合同的执行问题。但我们在这里感兴趣的是这样一种情况：对一些不常发生的事件，公司没有具体规定被监督者应该采取什么样的应对措施，并且/或者没有设计考虑到该类事件发生的激励机制。许多这类事件相对并不重要：流水线上的工人观察到一个零件有一个不常见的瑕疵。他必须做出判断，这个零件是应该被丢弃还是应该被使用。可能存在如下激励，使他试图做出正确的判断：[35]如果他这样做了，他将可能被提升。但即使他没有动力这样做，而是随机地做出决定，公司对他的行为进行干预，但公司也不能获得好处。因为监督者来到流水作业线，检查零件并做出决定所花费的时间不值得监督者采取干预行为。但也有另外一些非常规的事件的发生是很重要的，此时干预就是非常必要的了。

成本

我们需要区分各种组织形式的不同种类的成本：（a）建立制度的固定成本。这主要同设计合适的规则和激励机制有关。（b）培训工人的成本，培训他们如何应对各种突发事件（如果工人们都能够做出最优选择，那么他们所需要了解的就只是全部的激励机制，但当工人们有很多不同种类的行动时，激励机制可能非常复杂）。据推测，教工人扮演信息过滤器的角色——也就是识别某类的意外情况并向监督者咨询——会比教他们既识别意外情况又采取正确的措施更节约成本。（c）当监督者需要干预工人的行为

34 例如，参见 March and Simon［参考文献 12］。

35 本文的分析和团队理论最主要的区别在于工人和经理（公司的所有者）的目标并不一定一样。其中不仅存在信息的有效传输问题，还存在——就如我们在前面的讨论中已经注意到的，当工人知道其工作的难度——如何诱使工人披露信息的问题。我们前面的讨论也列举了一个诱使工人向公司披露信息的（次优的）办法，使公司可以利用信息提供一个“更好的”激励机制。

时，不间断监督的成本。（d）使监督者同时担任咨询师角色的成本。（e）当意外事件发生的时候导致的可变成本。因此在第（3）类组织中，监督者同时扮演咨询师的角色，成本既包括工人向监督者咨询的成本，也包括监督者向工人传达命令的成本，而在第（4）类组织中，只有监督者向工人传达决定的成本。第一类成本是一次性的成本。第二类成本是每单位时间的成本，但成本的大小取决于工人流动率。第（c）类和第（d）类成本是每单位时间的固定成本，但其大小是可以变化的，取决于监督者提供服务的质量。最后，第（e）类成本是可变成本。因此，在一定程度上，对组织形式的选择可以视作是对固定成本和可变成本的权衡。

规则化的成本很可能是同规则带来的收益无关的，因此只有对期望收益——也就是收益值乘以意外事件发生的期望频率——大的事件采取这样的措施，支付那么多的固定成本才合算。选择雇佣监督者当咨询师还是让监督者监控状况，并适时进行干预，这取决于固定成本（一种是培训成本，另一种是监督成本）和期望的可变成本（这取决于意外事件发生的频率乘以意外事件的发生导致的可变成本的变化）的相对大小。对不经常发生的事件，其期望可变成本也低，采取干预的形式似乎是最优的。对于相对不经常发生的意外事件，与其花费大量的培训成本让工人来识别这种意外事件，不如安排专人来识别和传递信息。[36]

---

36 正式的证明如下：用 B_i 表示第 i 种组织形式下的总收益，R_i 是常规化的固定成本（对组织形式 1 和组织形式 2），t_i 是第 i 种组织形式下培训工人的成本，S_i/π 是监督成本，v 是由监督者向工人传达决策的成本，u 是由工人向监督者传递信息的成本，N 为工人的数量，π 是事件发生的频率，ρ（N）为事件发生后监督者发现该事件的概率。因此，为了决定采取哪种组织形式，我们比较：

$$\pi B_1 - R_1 - Nt_1 - S_1\pi,$$
$$\pi B_2 - R_2 - Nt_2 - S_2\pi,$$
$$\pi B_3 - Nt_3 - \pi(u+v) - S_3\pi,$$
$$\rho\pi(B_4 - v) + (1-\rho)\pi B_5 - S_4\pi,$$
$$\pi B_5$$

正常情况下，我们可以认为：

$$R_1 > R_2, S_1 < S_2 < S_3 < S_4,$$
$$B_5 < B_1 < B_2 \leqslant B_3 = B_4,$$

并且有

$$t_3 < t_1 \text{ 和 } t_3 < t_2。$$

从这里可以非常直观地看出，对组织形式的选择取决于各种参数。

这种观点看起来就似是而非的，并且当信息的数量是可变的，就没那么有说服力了：为什么不让众多的工人获得较少量的信息，而要使少数人拥有大量的信息？上面分析中所建议的在经济活动中形成专业化是需要信息的非凸性的。事实上，信息有两种非凸性。第一，一定量信息的价值是随着它被使用的频率的增加而增加的，即即使雇员和监督者获得信息的初始成本是一样的，在等级结构中，监督者会更频繁地使用该信息。一般商品不同，信息的使用不会减少一个人所拥有的信息"存量"——甚至在一定程度上由于是一个学习的过程，反而会增加更多的信息。虽然这给专业化提供了一定的理由，但工人可能容易获得某一类的信息，这就抵消了专业化带来的好处，因为获得某一类的信息是工人在执行其任务时的自然产物（例如，观察到他们所用原材料的瑕疵）。[37]另一方面，在监督者参加其他活动时，有许多种类的信息（例如，同市场现象有关的信息）也能够比较容易地被获得；此外，如果信息会影响到许多工人的工作，则由监督者来获得该信息并将其传递给工人会比由每个工人自己来获得该信息更节约成本。

其次，我们可以证明，一般而言，只获得一点点信息是不合算的：太少量的信息的价值是负的，也就是，成本往往要大于收益。[38]因此只对工人进行很少的监督是不合算的，要么就完全不监督，要么就要达到一定程度的监督。我想这可能不仅为定义工人在工作中自主权的大小给出了一个分析框架，也为定义公司的边界给出了分析框架。[39]

组织理论[40]的文献很久以前就意识到了获得和散布信息的成本，监督者在这些过程中的作用，以及由于信息的成本导致的理性的界限。我们在这节中试图要做的是将监督和信息技术的结构联系起来，说明是信息的基本非凸性导致了职能上的专业化，而等级关系就体现了这种职能的专业化。[41]

四、结论

本篇文章中分析的问题，可以被视作是我们已经在"甄别理论"的题

---

37 关于分权的许多文献都讨论过这种情况：信息自然地从外围出现，然后流向中心。

38 参见 Radner and Stiglitz [参考文献 19]。

39 独立公司的一个特点是当公司不能支付其债务时，就会自动地产生唯一的一种干预行为。当两个公司合并以后，需要母公司不断地投入一定的成本来监督被合并公司的经营状况。

40 例如，参见 March and Simon [参考文献 12]，以及 Williamson [参考文献 34]。

41 当然，这里的证明只是一种定性地分析，即我们可以证明在一些非常一般的条件下，信息存在着一些非凸性。要解释专业化需要这些非凸性在某种意义上是非常显著的。

目下讨论过的一类问题的一些特殊情况。甄别理论的核心问题是经济体中一些决策人（保险公司、政府、雇主）无法在不付出成本的情况下辨别出不同的或潜在不同的目标对象。

这里，我们关心的问题是如何辨别不同工人的努力程度和能力大小，以及如何辨别不同工作的难度。

当然，并不是所有的不完全信息都可以放在甄别理论的框架下进行分析，而事实上，第三部分就分析了一些不能把不完全信息问题视作是甄别问题的情况。

文章的大部分内容关注的都是公司（和它们的工人）是如何遇到同甄别有关的信息不完全问题的。甄别理论已经发现了经济体在应对这种信息的缺乏时所采取的措施：

（1）公司会试图降低信息的缺乏所带来的成本。保险单中的共付条款使得保险公司不用太多地关注投保人的行为，这里的奖金机制使得公司不用太多关注工人的努力程度。

（2）公司会试图间接地获得信息，通过工人的行为来推断信息。这方面的例子在第二部分中已经给出了，即工人的表现会揭示出关于工作的难度和个人能力的信息。

因此，第一部分的大部分可以被看作是在探讨应对不完全信息情况的这两种不同方法。另一方面，第二部分关注的是第二种方法。

（3）识别出目标对象的不同，会导致直接的花费，这类似于教育中“通过考试甄别”。这里我们把直接的考核交给监督者完成，由他来观察工人的努力水平或个人的能力。

这篇论文关注了对监督的需求和等级生产结构带来的回报。我们已经证明了不了解自己完成某个工作的能力的工人（或者面临其他风险的工人）愿意为监督付费，以便确保自己和其他工人比在直接的奖金制下还更努力地工作；而了解自己能力大小的工人[42]愿意为监督付费，以使他们的能力被辨识出来，从而获得能力租金；对于了解自己能力但面临其他能影响自己边际生产率的风险的工人，等级生产结构也能带来收益。

当然，使用等级生产结构是否能带来好处还取决于监督的收益和成本，而在这篇文章中我们忽略了监督的成本。要解释等级生产结构的演变过程的理论，必须要关注技术的变化是如何影响到这种生产组织方式的成本和收益

42 如果他们相信他们的能力要高于平均能力水平。对更愿意努力工作的工人也可以得到类似的结论。

的。这些都是需要我们再进一步探讨的问题。

参考文献

1. AKERLOG, G. , "A Theory of Information and Labor Markets," *Quarterly Journal of Economics* (forthcoming, 1975) .
2. ALCHIAN, A. , AND DEMSET, H. "Production, Information Costs, and Economic Organization," *The American Economic Review*, Vol. 62, No. 5 (December 1972) .
3. ANAND, S. , "An Optimal Earning Schedule," Oxford, mimeo.
4. ARROW, K. J. "Higher Education as a Filter," *Journal of Public Economics*, Vol. 2 (July 1973) .
5. ______, *The Limits of Organization*, New York: W. W. Norton, 1974.
6. ATKINSON, A. AND STIGLITZ, J. , "Alternative Approaches to the Redistribution of Income," Stanford mimeo, 1975.
7. CYER, R. M. AND MARCH, J. G. , *A Behavioral Theory of the Firm*, Englewood Cliffs, N. J. : Prentice Hall, 1963.
8. DOERINGER, P. AND PIORE, M. , *Internal Labor Markets and Manpower Analysis*, Lexingtong, Mass,: D. C. Heath, 1971.
9. HARRIS, J. E. , WACHTER, M. L. , AND WILLIAMSON, O. E. "Understanding the Employment Relation: The Analysis of Idiosyncratic Exchange," *The Bell Journal of Economics*, Vol. 6, No. 1 (Spring 1975) .
10. LERNER, A. P. , *The Economics of Control*, New York: 1944.
11. LUPTON, T. , ED. , *Payment Systems*, Penguin Modern Management Readings, 1972.
12. MARCH, J. G. AND SIMON, H. A. , *Organizations*, New York: John Wiley, 1958.
13. MARGLIN, S. , "What Do Bosses Do?", Harvard University mimeo.
14. MARRLEES, J. A. , "An Exploration in the Theory of Optimum Income Taxation," *Review of Economic Studies*, 1971.
15. ______, "Notes on Welfare Economics, Information and Uncertainty," in Balch, McFadden and Wu, eds. , *Essays on Economic Behavior under Uncertainty*, Amsterdam: North Holland , 1974.
16. ______, "The Optimal Structure of Incentives and Authority Within an Organization," Oxford, mimeo.
17. PENCAVEL, J. H. , "An Essay on the Economics of Work Effort and Wage Payments Systems," Stanford mimeo, July, 1974.
18. PIGOU, A. C. , *The Economics of Welfare*, London: MacMillan, 1920, Chapter VII.
19. RADNER, R. AND STIGLITZ, J. , "A Nonconcavity in the Value of Information." Stanford mimeo, 1975.
20. ROTHSCHILD, M. AND STIGLITZ, J. E. , "Equilibrium in Competitive Insurance Market: the Economics of Markets with Imperfect Information," *Quarterly Journal of Economics* (forthcoming 1975) .
21. SALOP, J. K. AND SALOP, S. C. "Self Selection and Turnover in the Labor Market," Paper presented to Econometric Society Meetings, Toronto, 1972.
22. SAMUELSON, P. A. "A. P. Lerner at sixty," *Review of Economic Studies*, Vol. 31 (June

1964), pp. 169 – 178.
23. SPENCE, A. M. "Competitive and Optimal Responses to Signal: An Analysis of Efficiency and Distribution," *Journal of Economic Theory*, Vol. 7 (March 1974) .
24. ______, "Job Market Signalling," *Quarterly Journal of Economics*, Vol. 81 (August 1973), pp. 355 – 379.
25. ______, *Market Stgnalling*, Cambridge: Harverd Univ. Press, 1974.
26. ______, AND ZECKHAUSEN, R., "Insurance, Information, and Individual Action," *The American Economic Review*, Vol. 61, No. 2 (May 1971), pp. 380 – 387.
27. STIGLITZ, J. E., "An Economic Analysis of Labor Turnover," Stanford Working Paper, 1975.
28. ______, "Equilibrium Wage Distributions," Cowles Foundation, 1974.
29. ______, "Information and Economic Analysis," Paper presented to AUTE Conference, Manchester and published in Aute Proceedings, M. Parkin, ed.
30. ______, "Labor Turnover and Wage Structure," Yale mimeo.
31. ______, "A Market with Heterogeneous Qualities," Stanford mimeo.
32. ______, "Risk Sharing and Incentives in Sharecropping;" *Review of Economic Studies* (April 1974) .
33. ______, "The Theory of Screening, Education, and the Distribution of Income," *The American Economic Review* (forthcoming 1975) .
34. WILLIAMSON, O. E., *Corporate Control and Business Behavior*, Englewood Cliffs, N. J.: Prentice Hall, 1970.
35. WILSON, R., "On the Efficient Scale of a Firm," Stanford University, August 1974.

奖金与激励：关于薪酬与竞争的一般理论*

本文研究了在信息不完美条件下竞争性薪酬安排的作用（其中薪酬依赖于个体的相对绩效）。竞争性薪酬安排在风险性、激励性以及灵活性方面都具有令人满意的性质；在经济环境发生变化的情况下，竞争性薪酬安排能够依据变化自动地对薪酬以及激励做出相应的调整。当经济环境发生了巨大的改变时，可以证明竞争性薪酬安排要优于那些仅仅依赖个体绝对绩效的薪酬安排；在极端的情况下，当参与竞争的人数变得非常大时，有可能使得个人效用逼近一级最优的水平（信息完美的情况）。我们要研究竞争机制的设计，包括如何最优化奖赏和惩罚的水平，以及如何优化绩效评价标准的绝对风险和相对风险水平。这里的分析将同样可以视为对于单个委托人多个代理人问题的一个研究。

一、引言

现代资本主义经济的一个主导特征就是竞争在经济中发挥着重要的作用，正是由于竞争才确保了工作激励的存在。在竞争性的市场中，企业内部或者企业之间很少是通过计件工资率（在传统的竞争模型里它等于边际产品价值）来结算薪酬的。在更常见的情况下，至少有一部分薪酬是基于个体的相对绩效的：分数最高的学生才给 A 的成绩，业绩最好的推销员才得到很高的红利，市场占有率最高的企业才成为市场的领导者，而只有最好的经理人才有可能晋升为企业副总裁。相比阿罗—德布鲁一般均衡模型里单一的静态价格竞争模式，专利竞赛和体育竞技可以更好地反映出竞争概念的含义。

---

* Barry J. Nalebuff and Joseph E. Stiglitz, Prizes and Incentives: Towards A General Theory of Compensation and Competition，写作本文时，前一位作者属哈佛大学，后一位作者属普林斯顿大学。本文曾经作为 1980 年 7 月在柏林国际管理学院召开的国际企业组织大会的会议论文。与 Sherwin Rosen, Jim Mirrlees, Terence Gorman, Richard Zeckhauser, Kevin Roberts, Ed Lazear, Felix Fitzroy, Partha Dasgupta, Joe Farrell, Jerry Green, Jim Poterba, Roger Koenker, Bengt Holmstrom, Steve Salop, 以及 Hal Varian 的讨论令我们受益匪浅。十分感谢在柏林国际管理学院的大会上以及在威斯康星大学、密执安大学、芝加哥大学、宾夕法尼亚大学、联邦贸易委员会以及伦敦经济学院的讨论会上与会者对于本文所做的评论。我们要特别感谢 Alvin Klevorick 以及两位匿名审稿人提出的改进建议。非常感激国家科学基金会以及罗德斯奖学基金会对本文的资助。

过去的两年以来，竞争作为一种激励手段的作用让人们重新燃起了对它的兴趣。早期莱希尔和罗森（Lazear and Rosen，1981）以及斯蒂格利茨（1980）的研究，后来被格林和斯托基（Green and Stokey，1981）、哈特（Hart，1981）、霍姆斯特姆（Holmstrom，1982）、穆科基（Mookherjee，1981）、纳尔巴夫（Nalebuff，1982）、费茨罗伊（Fitzroy，1981）以及奥基菲、维斯库斯和泽克豪斯（O'Keefe，Viscusi and Zeckhauser，1982）大大地扩展了。最近的这些文献已经开始着重描述如下的情况：基于个体相对绩效的薪酬安排要优于基于个体绝对绩效的薪酬安排。

我们关注的这个问题之所以会产生，是由于工人（经理人）的投入（努力程度）是无法用直接的和无成本的方式观察到的。因此企业只能通过努力监督投入，或者只能通过设计有效的薪酬安排使得报酬作为变量的函数（最广义的定义包括晋升，养老金，等等）具有如下性质：（a）本身就是投入的函数；（b）相比投入的观测成本要小。

任何薪酬安排都具有以下三个关键特征：

（a）*风险*。将薪酬与产出相联系的激励方案不可避免地使得工人们承担了风险；如果工人们是风险规避的话，那么结果就是福利的损失。

（b）*激励水平*。为了缓解风险问题，在绝大部分个体的薪酬安排中，都有相当的部分采取不直接与产出挂钩的形式。

（c）*灵活性*。在某种情况下是“正确”竞争性的薪酬方案，在另一种情况下也许就不是了。原则上，（合同里）对应于每一组环境变量都应该存在着不同的激励安排。然而，在现实中，起草那样复杂的合同显然成本太高；而且对于所有的合同签订方而言，对许多相关的环境变量的观察都不是没有成本的。这样，某一种激励安排必须能够广泛地适应于不同的环境。缺乏灵活性被普遍地认为是计件工资率的致命缺陷：实行计件工资率是有成本的，而且其结果也是有争议的。

正是由于竞争性的薪酬安排具有灵活性，这使得它非常具有吸引力。如果对于某个人而言某种工作十分简单，那么很可能对于他的竞争对手们而言也是十分简单的。考察一下当员工的报酬与其相对绩效挂钩的时候所产生的技术创新效应，创新会使得个人的产出超越对手变得简单，从而在这个新的均衡下每个人的产出都增加了，但报酬并未改变。这样每单位的报酬就自动地降低了。相应地，尽管很难对于一个难度水平未知的考试设计出一套定值的评分系统（$90=A$，$80=B$），但是却可以通过对学生成绩的相对水平来评分，从而，使得一套评分体系可以适用于不同的课程、不同的班级以及不同

能力的群体。

在许多情况下，个人（或企业）的绩效都传递了有关环境的信息，而这些信息都完全可以而且应该纳入到薪酬安排之中（Hart，1981；Nalebuff and Stiglitz，1983；Stiglitz，1982）。竞赛就是一种有效的方法；另一种方法就是将个人的报酬与绩效评价标准挂钩，而绩效评价标准又反过来与团队的绩效挂钩。

（竞争性）薪酬安排还有一个重要的功能：它能够甄别不同能力的员工，并且予以区别对待（Stiglitz，1975）。本文着重探讨了薪酬安排的激励性质，因此，除非下文中有提及，我们都假设所有人的能力是相同的。

我们将首先在第三部分到第五部分中给出薪酬安排的一般理论。然后再考察竞赛理论（第六部分至第八部分），并且将其与其他的薪酬安排相比较（第九部分至第十一部分）。第十二部分至第十四部分是对该基础理论的重要的扩展。

二、结论概述

在定义了一般性的问题之后，我们把注意力集中到几个特例之上：比如我们的大多数分析都假设了产出是努力程度的线性函数。我们考察了在何种条件下竞争性的薪酬安排（其中报酬仅仅依赖于个人的相对产出排名）要优于基于个人绝对产出的薪酬安排。由于分析过程相当繁琐，把主要的结论先列出来也许是一个好办法。

（1）求出次优解一般要求知道每个代理人的产出。如果产出能够通过恰当地使用统计量得到刻画的话，这个问题就能够得到简化。例如，在当代理人的数目足够大的情况下，最优的报酬可能仅仅依赖于个人产出和平均产出。

（2）如果代理人是风险中性的，那么通过恰当地设计竞赛可以达到最优的均衡。

（3）在风险很小从而奖金不太丰厚的情况下，不存在纯策略纳什均衡。

（4）尽管很难对努力程度进行监督，从而也很难调动员工的积极性，会产生普遍的委托—代理问题，但是其次优均衡解所达到的努力程度有可能会大于最优解的努力程度。即便如此，这只是一个道德风险的问题：努力工作可以提高个人的效用。

（5）将竞赛作为激励手段可以促使代理人放弃他们本来的风险厌恶倾向，从而变得更加愿意冒险，更敢于采用利润更高的生产技术。

(6) 当参赛者人数众多时：(i) 恰当地设计竞赛经常可以达到最优的结果；(ii) 惩罚绩效最差的个体要比奖励绩效最佳的个体更能调动员工的积极性。

(7) 竞赛机制可能要优于基于个人绝对绩效的薪酬安排，特别是当与一般的环境变量相关的风险非常大的时候。

(8) 即使代理人之间是独立的，两人竞赛的结果也可能绝对占优于基于个人绝对绩效的线性报酬方案。

(9) 通常通过要求参赛者必须与对手拉开足够大的差距，才可以获得锦标赛胜利的方法可以改进简单竞赛的结果。

(10) 竞赛中不存在纯策略均衡，当在只有两个参赛者的情况下，混合策略均衡的结果可以相对接近最优的结果。不过，随着参赛者的人数越来越多，竞争变得越来越激烈，最终所有的剩余都散尽了。

(11) 有多重奖项的锦标赛一般是不能够替代非线性薪酬方案的（反之亦然）。运用两种经常使用的误差分布的例子，我们证明了即使可以使用 n 种不同的奖金，最优解至多也只需要三个不同的奖金。

(12) 当产出与努力程度不是线性关系的时候，锦标赛机制可能或不可能实现一级最优努力水平的边际激励。每一种可能的情况我们都给出了例子。

三、基本模型

考虑一种非常典型的经济环境。其中代理人 i 的产出 Q_i 是可以观察到的，并且产出是他的努力程度 μ_i 的随机函数：

$$Q_i = Q(\mu_i, \theta, \epsilon_i) \tag{1}$$

随机性来自于公共的“环境”变量 θ，以及某个“私人”的扰动 ε_i。对于式 (1) 可能会有几种不同的理解方式。例如说，在农业经济条件下，θ 代表了某地区的整体天气特征，而 ϵ_i 代表了该地区某个农场的天气情况（下雨）。或者 θ 代表了某一组任务的整体难度，而 ϵ_i 代表了个人在执行该组各项任务时的比较优势（或者比较劣势）。分析过程中值得注意的是：

(a) $Q_{\mu\theta} \neq 0$，自然状态会影响到努力程度的回报；而且

(b) 对于每一个 θ 值，如果知道了 Q_i 和 θ 的值，就可以准确地推断出 μ 的值。否则，就会有：

(a) 自然状态不会影响到努力程度的回报；或者

(b) 通过观察到产出可以直接监督员工的努力程度。

在全篇文章中，我们都假设合同签订前 θ 和 ϵ_i 是未知的；签订后工人

观察到了 θ 并且决策出 μ 的值，这些决策过程企业是无法观察到的。关于 θ 和 ϵ 的概率分布企业和工人都是知道的。我们将分析重点集中于这个特别的情况：

$$Q_i = \mu_i\theta + \epsilon_i \tag{1a}$$

其中 G 是 ϵ_i 的分布函数，g 是其分布密度。并且 $E[\epsilon_i] = 0$，$E[\epsilon_i\epsilon_j] = 0, E[\theta] = 1$。采取线性的生产函数是为了分析的简便。[1] 放松这个条件的效果将在第十三部分予以讨论。在相关的委托—代理文献之中，式（1a）是十分重要的，因为莫里斯（Mirrlees，1975）的最优收入税收模型可以视为此模型在 $\sigma_\epsilon^2 = 0$ 条件下的一个特例；而范里安（Varian，1980）关于社会最优保险的分析也可以看作在 $\sigma_\theta^2 = 0$ 条件下的一个特例。

我们的分析主要关注于完全竞争市场，其中企业是风险中性的而且期望利润也削减至零。期望收入等于期望产出的薪酬安排被称为是可行的。在多个代理人的情况下，这就意味着产出向量 $\{Q_1, \ldots, Q_n\}$ 的值就是 $\sum Q_i$。在专利竞赛的模型里，$1/Q_i$ 可以代表第 i 种专利的开发时间，不过此时更合理的假设应该是只有最大值 $\max\{Q_i\}$ 才影响决策，而且这样可行的定义也会不同了（Nalebuff and Varian，1981）。

假设代理人是同质的，其效用函数对于收入 Y 和劳动 μ 是加性可分（additively seperable）的。当收入是个随机变量，期望效用 W 为，

$$W = E[U(Y) - V(\mu(\theta))] \tag{2}$$

其中收入的边际效用为正而且是非递增的（个人不是偏好风险的）。而努力程度产生负面效用是正的而且还是递增的：

$$U'(Y) > 0, U''(Y) \leq 0, V'(\mu) > 0 \text{ 和 } V''(\mu) > 0 \tag{3}$$

四、充分统计量

由于只有产出集合 $\{Q_i\}$ 是可以观察到的，从而薪酬安排仅仅是 $\{Q_i\}$ 的函数

$$Y_i = Y(Q_1, \cdots, Q_i, \cdots, Q_n) \tag{4}$$

为了使得薪酬安排具有更大的灵活性，委托人必须设计出一个两阶段的包含信号的博弈。恰当的激励结构可以使得个体在宣布 θ 的时候发送出相关信

1　当然，这并不是最一般的线性函数形式；我们可以把式（1b）写为 $Q_i = \mu_i\theta + \epsilon_i + \gamma + \mu_i\eta_i$，即除了存在着一般性的总产出效应之外还存在着特别的边际产出效应。

息。此类揭露机制在纳尔巴夫和斯蒂格利茨的论文（1983）中有所讨论。

很明显，一般的相对绩效的薪酬安排应该不比基于个人的绝对绩效安排差：

$$Y_i = Y_i(Q_i) \tag{5}$$

问题就在于究竟什么时候前者会优于后者？有时候委托人别无选择，只能将所有的报酬都与某一单一产出挂钩。在专利竞赛的例子中，可能仅仅只是存在一个可以观察的产出变量——开发时间。失败者可以修改目标，从头再来，但是一般而言不会产生可以观测到的产出。

采取多重产出来安排每个代理人薪酬的好处就在于，通过观察整个产出集合 Q_i 往往可以收集到一些 θ 的信息。而这些信息可以用来推断出 μ 的值，从而使得委托人能够紧缩报酬安排并且获利。

定理 1：如果所有产出是独立的（$\sigma_\theta^2=0$），那么通过观察全部的产出 Q_i 不会获得任何有关代理人的努力水平值 μ_i 的更多信息。在这种情况下，最优薪酬安排是基于个人绝对产出的安排。

证明：这是一个简单而又有效的证明方法。取 Y_i（Q_i）使得满足下式：

$$U(Y_i(\tilde{Q})) = E[U(Y_i(Q_1,\cdots,Q_i,\cdots,Q_n \mid Q_i = \tilde{Q}))] \tag{6}$$

假设所有工人的期望效用作为其产出的函数是相同的。由于他是风险中性的，所以有下式成立：$Y_i(\tilde{Q}) < E[Y(Q_1,\cdots,Q_i,\cdots,Q_n \mid Q_i = \tilde{Q})]$。委托人可以对所有的代理人提供完全相同的激励而且期望获得更多的利润。当代理人是独立的情况下，基于个人绝对产出的薪酬安排更加有效只不过是充分统计量定理的一个特例罢了。通常是有可能为代理人预期中的公共元素 θ 找到一个充分的统计量 T（Q）的，这样最优薪酬安排就能够简化为

$$Y_i = Y_i(Q_i, T(Q)) \tag{7}$$

从直觉上讲，所有有关个人 i 的信息都包含于 Q_i 和 T（Q）之中的。使用任何其他的变量都会向薪酬安排中引人随机扰动从而不是最优解。[2] 充分统计量定理的一个应用就是前面结果中的定理 1。如果 $\sigma_\theta^2=0$，那么 θ 的一个充分统计量就是常数，于是最优的薪酬安排就是基于个人绝对产出的薪酬安排。其他的应用也是很直接的。如果 θ 和 ϵ 是服从正态分布的随机变量（其方差已知），那么 Q 的均值 $\bar{Q}$ 就是 $\theta\mu$（θ）的一个充分统计量，因此也就是 θ 的一个充分统计量。例如在一个班级中，当学生的人数很多的时候，平均

2　我们十分感谢 Joe Farrell 对此问题的深入地探讨。类似的关于充分统计量的研究成果和证明请参见 Farrell（1981），Grossman and Hart（1980），Holmstrom（1982）以及 Arrow（1970）。

分数是能够揭示测试的难度的。当参赛者的人数越来越多，由 $\overline{Q} = \theta\mu(\theta)$ 导出的 θ 的统计量会逐渐收敛到 θ 的真实值。在极限时就可以得到 θ 的充分统计量，因为那时我们就知道 θ 的真实值了。然后可以运用充分统计量定理来证明次优的薪酬安排仅仅依赖于个人绝对产出以及 θ，即 $Y_i = Y_i(\theta, Q_i)$。即使员工的人数有限，将薪酬安排与个人产出相对平均产出的大小相挂钩也是更好的安排，因为它可以减少与 θ 相关的扰动（就如同第十一部分里相对绩效安排所做的那样）。

我们可以把前面提出的薪酬安排问题视为一个统计量的问题。企业无法知道工人的努力水平。它试图通过了解博弈的结构以及观察工人的行为以做出“最优的”预测，然后在此基础上决定最优的薪酬安排。这种方法的一个困难之处在于所用的统计量同时具有风险和激励效应，而且这两者必须在抉择最优统计量的时候都予以考虑。

五、最优安排：一个基准

在以下的论述中，有必要将待选的薪酬安排同在 θ 和 μ 是可以观察的完美信息条件下的均衡情况比较一下。在风险中性的企业和风险厌恶的工人之间达成的最优薪酬安排，工人们的风险被全部保险了，得到预定的报酬 $\overline{Y}$，而预定报酬 $\overline{Y}$ 仅仅依赖于期望产出。当收入的边际效用乘以单位努力程度的产出正好等于努力的边际效用损失的时候，努力的水平就决定了：

$$\theta U'(\overline{Y}) = V'(\mu(\theta)) \rightarrow \mu^*(\theta) \equiv V'^{-1}(\theta U'(\overline{Y})) \tag{8}$$

于是有

$$\overline{Y} = E[\theta\mu(\theta)] = E[\theta V'^{-1}(\theta U'(\overline{Y}))] \tag{9}$$

六、竞赛的基本结构

名次锦标赛制是指这样一种薪酬安排，参赛者的报酬依赖于他们在结果中的相对名次，而不是产出的绝对大小。当只有两个参赛者时，优胜者将获得优胜工资 Y_W 而失败者得到败者工资 Y_L。不同于标准的边际分析，这里优胜者的产出不一定值 Y_W 的工资。付给优胜者的薪水要高于优胜者的边际产出，这就等于抛出了“胡萝卜”来调动参赛者更大的积极性。

委托人在完全竞争市场中期望利润为零的条件下，寻求员工效用的最大化。如此，$Y_W + Y_L$ 就必须等于期望产出的总和 $2\overline{Q}$（这其实是委托人在代理人效用约束条件下利润最大化问题的对偶问题）。这样改变一下记号会更方便。令

$$\bar{Y}=E[Q_i]=\frac{Y_W+Y_L}{2} \text{和} x=\frac{Y_W-Y_L}{2} \tag{10}$$

任何两种奖金都可以视为某个固定收入 $\bar{Y}$，加上风险（或者奖金）x。

个人努力。个人 期望效用是其获胜概率的函数。后者又反过来取决于个人自身努力程度以及对手的努力程度，环境变量 θ 和 ϵ 的分布，获胜的概率可以表示为 P（μ_1，μ_2，θ）。某个代理人的期望效用可以表示为

$$W=PU(\bar{Y}+x)+(1-P)U(\bar{Y}-x)-E[V(\mu)] \tag{11}$$

通过将报酬与竞赛名次挂钩而不是与努力程度挂钩，我们引入了随机扰动项 ϵ，ϵ 的期望为零，与努力程度 μ 不相关。我们用一个与努力程度 μ 相关的扰动项 x 替代 ϵ。如果参赛者努力工作，那么他获胜的机会就大，这样 x 为正的概率 P 就上升了。在对称均衡之中，获胜的概率是1/2，这样期望的奖金（x 或 $-x$）就是零。

当工人努力工作的边际效用损失正好等于努力工作增加的获胜机会的价值时，他的努力水平就决定了。当观察到 θ 值之后 ，努力水平就由下式决定：

$$[\partial P(\mu_1,\mu_2,\theta)/\partial\mu_1]\Delta U-V'(\mu(\theta))=0 \tag{12}$$

其中

$$\Delta U=U(\bar{Y}+x)-U(\bar{Y}-x)$$

由于 μ_1 为 μ_2 的函数，可对上述等式求解，同样地可以得到 μ_2 是 μ_1 函数的一个对称的反应函数。[3] 尽管很明显有可能存在不对称均衡，也就是说其中一人努力工作，另一人偷懒。这里我们把注意力放在 $\mu_1=\mu_2$ 的对称均衡上。

努力的回报。在两个参赛者情况下，产出大的一方获胜。如果努力水平为 μ_1 的第一个代理想要击败第二个代理人，那么 ϵ 的实际值必须满足下式

$$\theta\mu_1+\epsilon_1>\theta\mu_2+\epsilon_2 \tag{13}$$

给定 ϵ_2，此事件发生的概率为

$$1-G(\theta(\mu_2-\mu_1)+\epsilon_2) \tag{14}$$

为了计算获胜的总概率，对于所有 ϵ_2 的可能值作积分，其权重取决于 ϵ_2 的密度 g（ϵ_2）。因此

---

3 如果 μ_2 是 μ_1 的连续递增函数，并且如果对于很大的 μ 值 ，有 $d\mu_2/d\mu_1<1$（如果某个人努力工作，那么另一个人也会努力工作，但是后者的努力程度的增加量不如前者），那么就存在着一个（对称的）均衡。然而，第二个人也是有可能降低努力程度来应对第一个人提高努力的状况。而且实际上，反应函数可能不是连续的。本例中可能不存在对称均衡。但是可能存在着非对称均衡，或者不存在纯策略均衡，抑或存在着多重均衡（参见附录部分的讨论）。

$$P(\mu_1,\mu_2,\theta) = \int[1 - G(\theta(\mu_2 - \mu_1) + \epsilon_2)]g(\epsilon_2)d\epsilon_2 \quad (15)$$

对称解要求 $\mu_1 = \mu_2$，解得 $P = 1/2$，第一个代理人通过更加努力获胜的概率为

$$\partial P(\mu,\mu,\theta)/\partial\mu_1 = \theta\int g(\epsilon_2)g(\epsilon_2)d\epsilon_2 = \theta\bar{g} \quad (16)$$

其中 $\bar{g} = E[g(\epsilon)]$。[4]

竞赛的里程碑方程。将式（16）代入一阶条件式（12）中得到表现竞赛均衡的基本方程：

$$\theta\Delta U\bar{g} = V'(\mu(\theta)) \quad (17)$$

努力工作的效用损失正好等于因为努力工作使得获胜的机会增加带来的效用收益，预期的酬劳必须等于预期的产出

$$\bar{Y} = E[\theta\mu(\theta)] = E[\theta V'^{-1}(\theta\Delta U\bar{g})] \quad (18)$$

最优竞赛是在可行性约束条件下满足参赛者预期效用最大化的组合（$\bar{Y}$，x）。

竞赛的核心性质。里程碑方程阐释了锦标赛制的一个重要性质：努力程度受 θ 的影响而期望收入则不然。此性质广泛适用于各种不同自然状态正是一个好的薪酬安排的重要特征。它甚至有可能达到一级最优情况下的努力程度，假设

$$U'(\bar{Y}) = \bar{g}\Delta U \quad (19)$$

这意味着

$$\theta U'(\bar{Y}) = \theta\Delta U\bar{g} = V'(\mu^*(\theta)) \quad (20)$$

尽管有可能达到最优情况下的努力程度，这个竞赛却无法达到同最优情况下一样的效用水平。在锦标赛制中，代理人必须承担风险，因为奖金要么为 x 要么为 $-x$。显然，倘若代理人是风险中性的，那么收入的这种变化就是不重要的，按式（19）制定奖金就将达到完全信息条件下最优的结果。一般而言，委托人会牺牲掉一部分效率来降低由代理人产生的风险。

定理2：更高的奖金会激发更大的努力，从而增加平均收入。

证明：为了阐明努力程度是如何随着奖金的变化而变化的，对工人的一

4　可以看到在很广泛的条件下二阶条件都可以满足。将式（12）微分，要求 $\partial^2 W/\partial\mu^2 = \theta\Delta U\int g(\epsilon)g'(\epsilon)d\epsilon - V''(\mu) < 0$。如果扰动是对称的，那么第一项为零，第二项由假定为负。更一般的情况下，如果 $\theta\Delta U[g^2(\bar{\epsilon}) - g^2(\underline{\epsilon})] < 2V''(\mu)$，二阶条件就会满足，其中 $\bar{\epsilon}$ 和 $\underline{\epsilon}$ 是 ϵ 的最大值和最小值。

阶条件式（17）求导可得：

$$\theta\bar{g}[U'(\bar{Y}+x)\{d\bar{Y}/dx+1\}-U'(\bar{Y}-x)\{d\bar{Y}/dx-1\}]=V''(\mu)d\mu(\theta)/dx \tag{21}$$

乘以 θ 再取预期函数，得到

$$d\bar{Y}/dx=\{\bar{g}SE[\theta^2/V''(\mu(\theta))]\}/\{1-\Delta U'\bar{g}E[\theta^2/V''(\mu(\theta))]\}>0 \tag{22}$$

其中

$$S=U'(\bar{Y}+x)+U'(\bar{Y}-x)\text{ 而且}$$
$$\Delta U'=U'(\bar{Y}+x)-U'(\bar{Y}-x)<0$$

尽管代理人都相信努力工作会增加获胜机会 P，但是实际上均衡时，两人的努力程度一样，因此 P 的值还是 1/2。但是由于两人的努力程度上升，新的可行的薪酬安排一定会增加等量的产出 $\bar{Y}$。

当奖金大到可以引致最优努力水平的时候，代理人可能会发现失败的那种自然状态是非常令人讨厌的。在这种情况下，$d\bar{Y}/dx$ 可能会大于 1，而竞赛也表现出某些“你死我活”的味道（Akerlof，1976）。随着奖金数额的增加，代理人努力工作的结果使得失败的自然状态下的收入变为 $\bar{Y}-x+(d\bar{Y}/dx-1)dx$。这远远大于原来的值。相对于奖金的增加，他额外付出的努力要更多一些。

最优赛制。现在我们能够解出只有两个参赛者的计名次锦标赛情况下最优赛奖金 x。回忆在对称均衡之中，$P=1/2$，而且努力水平的值由式（17）决定。因此由式（11）期望效用为，

$$W=(1/2)[U(\bar{Y}+x)+U(\bar{Y}-x)]-E[V(V'^{-1}(\theta\Delta U\bar{g}))] \tag{23}$$

再对 x 求微分，在最优奖金 $\tilde{x}$ 处有，

$$\begin{aligned} dW/dx &= (1/2)S\bar{Y}'+(1/2)\Delta U'-E[V'd\mu/dx] \\ &= (1/2)S\bar{Y}'+(1/2)\Delta U'-E[\bar{g}\Delta U\theta d\mu/dx] \\ &= [(1/2)S-\bar{g}\Delta U]\bar{Y}'+(1/2)\Delta U'=0 \end{aligned} \tag{24}$$

其中 $\bar{Y}'$ 和式（22）里的一样，而值 $V'=\theta\Delta U\bar{g}$ 则由式（17）得到。

定理 3：在次优情况下存在着努力水平供给不足。而可观察到的代理人的努力提高程度能够改进他的福利水平。

证明：在最优奖金处，如果委托人能够观察到努力水平的边际增加值，那么福利的变化值为

$$dW/d\mu=[(1/2)S-\bar{g}\Delta U]\bar{Y}'>0 \tag{25}$$

此式由式（24）可以直接推出。

七、采用竞赛后次优解的性质

本部分主要讨论在什么情况下某一套奖金体系有可能成为有效的薪酬安排。

非凸性产生的问题。一般而言，我们的奖金体系模型是存在着纯策略纳什均衡的。当然也有一个很重要的例外情况。

定理4：当ϵ的方差变得很小，最大化就会存在着非凸性问题，从而破坏了纯策略次优解的存在性。

证明：当σ_ϵ^2趋近于零，努力工作所增加的获胜机会$\theta\bar{g}$就会变得无穷大。由定义$\bar{g}=E[g(\epsilon)]$。当期望为零的随机变量方差趋近于零，那么其密度函数必定缩小为纵轴原点之上的某一点。从而密度函数的均值就变得任意大。例如，当$g(\epsilon)$为均值为零，方差为σ_ϵ^2的正态分布密度，那么$\bar{g}=1/[4\pi\sigma_\epsilon^2]^{1/2}$。

从直觉上讲，如果存在着很小的扰动，那么只需略微比对手多一点努力就能够确保参赛者获胜。随着$\bar{g}$变得很大，任意小的奖金额都可以保证代理人付出最优的努力水平。然而，如果奖金额很小，那么对称均衡解就只是一个局部最优解。为什么非要工作呢？如果代理人选择的努力水平不是μ^*而是0，那么他当然会输。然而由于奖金额如此之小，尽管他损失了获胜机会带来的价值ΔU，但是却节省了努力工作产生的效用损失，$V(\mu^*)-V(0)$。

不工作并不就是拒绝参加博弈。在合同上作弊与其类似。当扰动的方差变得任意小，毫无疑问，低产出是由不够努力工作而不是运气不好造成的。但是奖金支付体系过于简单因而很难阻止例外的出现。因为奖金额仅仅依赖于名次，所以在此情况下，即使是仅差一点点的运气，也可能功败垂成！如果对于不同的低产出有不同的低惩罚的话，我们就回到了一般的非线性薪酬安排之中。在斯蒂格利茨（1980）和纳尔巴夫（1982，chapt. 2）的论文中详细探讨了关于非凸性所引发的问题的各种解法。在接下来的讨论之中，我们把注意力主要放在有足够多的扰动条件下对称解的确构成了均衡的情况。

与最优的努力水平比较。概述中所说的次优均衡解所达到的努力水平，有可能会大于最优解的努力水平很容易招致怀疑。如果最优解无法达到是因为委托人无法监督代理人的努力程度，所以导致努力水平的供给不足，那么乍看起来，仅仅引入奖金体系就能够产生比最优解的努力水平还大的努力水

平是很令人怀疑的。然而引入奖金体系会使得代理人承担一些风险，而此风险所引起的成本与其收入是相关的。通过更加努力工作挣得更多的收入，绝对风险规避系数递减的代理人可以降低其风险规避的程度。在纳尔巴夫（1982）的例子里，他证明了此种效应可能大到足以使得次优均衡解所达到的努力水平大于最优解的努力水平。

这个结果对于效用函数的形式是很敏感的。在莱锡尔和罗森（Lazear and Rosen，1981）的框架下，期望效用被定义为：[5]

$$\widetilde{W} = E[U(Y - V(\mu))] \tag{26}$$

此处次优的努力水平一定永远小于最优努力水平的。这一结果由下面的事实直接推得：最优努力水平的决定 $V'(\mu^*) = \theta$ 是不受收入的随机波动影响的。

竞赛中 θ 的方差变化产生的效应。倘若奖金体系是最优的，θ 的方差变化会直接影响到期望效用，并且通过对努力水平来间接影响期望效用（而不是间接通过 x）。他对于 θ 的分布的影响是很难确定的，而且通常取决于努力的效用损失 $V(\mu(\theta))$，究竟是 θ 的凸函数还是凹函数。

定理5：对于二次型的努力效用函数形式，福利函数就是 θ 的方差的增函数。

证明：由于

$$V(\mu) = \mu^2/2 \Rightarrow \mu(\theta) = \theta\bar{g}\Delta U; \qquad \bar{Y} = \bar{g}\Delta UE[\theta^2];$$
$$E[V(\mu(\theta))] = (1/2)\bar{g}\Delta U\bar{Y} \tag{27}$$

并且

$$\partial\bar{Y}/\partial E[\theta^2] = \bar{g}\Delta U/\{1 - \bar{g}\Delta U'E[\theta^2]\} > 0 \tag{27a}$$

为了计算出 θ 的分布变化的总效应，注意到

$$W = (1/2)[U(\bar{Y} + x) + U(\bar{Y} - x)] - (1/2)\bar{g}\Delta U\bar{Y} \tag{28}$$

$$dW/dE[\theta^2] = \frac{dW}{dx}\frac{dx}{dE[\theta^2]} + \frac{\partial W}{\partial \bar{Y}}\frac{\partial \bar{Y}}{\partial E[\theta^2]} \tag{29}$$

$$= [(1/2)S - (1/2)\bar{g}\Delta U - (1/2)\bar{g}\Delta U'\bar{Y}]\partial\bar{Y}/\partial E[\theta^2] \tag{29a}$$

$$> [(1/2)S - \bar{g}\Delta U]\{\partial\bar{Y}/\partial E[\theta^2]\} > 0 \tag{29b}$$

其中 $dW/dx = 0$，而最后一个不等式可以由决定 $\bar{x}$ 的一阶条件式（24）得到。在此例中，θ 的期望保持扩展提高了效用水平。*于是，θ 方差为零所带来的福利损失是由竞赛所带来的福利损失的上限*。更一般地说，此结果依赖于关于 V'' 的假设，后者决定了新的分布是如何影响产出和工作的效用

5 可以这样来理解这个特殊的效用形式：工作难度是和收入负相关的。

损失的。

斯蒂格利茨（1980）证明了 θ 方差的增加可以改善在竞赛机制下福利状况，而降低计件工资制下的福利状况所需要的条件。从直觉上讲，由于 θ 是非常不稳定的，为了给计件工人足够的激励使其努力工作，计件工资率必须为正，而且工人必须承担所有的与 Q_i 大幅波动相关的风险。相比之下，竞赛机制可以达到最优的努力水平，而且相对于最优的福利损失的货币价值要严格小于由奖金产生的风险损失（是 $-x^2U''(\bar{Y})/2U'(\bar{Y})$ 的同阶项），而后者在 σ_θ^2 增加时保持不变。[6]

尽管风险常常是以 θ 的方差表示的，但是还可以适当考虑一些 θ 的高阶矩变化的影响。考虑一个 θ 的期望和方差不变而 θ^2 的方差递增的期望保持扩展。在努力水平的效用函数为二阶型的形式时，由式（27）和式（28）福利水平仅仅是 $E[\theta^2]$ 的函数，于是有，

$$dW/dE[\theta^4]=0 \tag{30}$$

当工人努力水平的效用函数为二阶型的形式时，θ 的分布对于竞赛影响仅仅通过 θ 的前两个矩起作用。

八、对标准锦标赛制进行改进

本部分将要介绍两种可以改进均衡结果的标准锦标赛制修改方案。

领先距离足够多才算获胜。任何委托—代理问题激励安排的目的都是为代理人努力工作提供足够大的激励而又尽量不让代理人承担风险。引入“领先距离”有可能降低所付奖金而又能够保持边际激励水平不变。不再仅仅以代理人在产出的相对位置来评判胜负，委托人只有在代理人的产出领先竞争对手足够多的时候才判断优胜。在赛马和奥运会游泳比赛中，哪怕仅仅领先一头或者一丝毫发也是冠军。然而在市场中，一个企业作为某个产业的领导者其市场份额必须要比其他的竞争者高出许多才行。

假设必要的领先距离是 γ 存在着三种可能性：赢、输或者平局。假设平局的奖金为 $\bar{Y}$ 而胜利的奖金如前所述为 ΔU。在两个参赛者的情况下，对称均衡中期望效用是，

$$W=(1-2P)U(\bar{Y})+PU(Y+x)+PU(\bar{Y}-x)-E[V(\mu)] \tag{31}$$

如果第一类代理人获胜必须有

6　结论还能够更加一般化，可以证明当 θ 的方差变得很大时，即使采用非线性计件工资制，竞赛机制要严格占优的成立条件。

$$\theta\mu_1 + \epsilon_1 > \theta\mu_2 + \epsilon_2 + \gamma \tag{32}$$

将式（15）和式（16）变形可以得到：

$$P(\mu_1,\mu_2,\theta,\gamma) = \int[1 - G(\theta(\mu_2 - \mu_1) + \epsilon_2 + \gamma)]g(\epsilon_2)d\epsilon_2 \tag{15a}$$

在对称均衡解处，第一类代理人努力工作所增加的获胜机会是

$$\partial P(\mu,\mu,\theta)/\partial\mu_1 = \theta\int g(\epsilon_2 + \gamma)g(\epsilon_2)d\epsilon_2 = \theta\bar{g}(\gamma) \tag{16a}$$

其中 $\bar{g}(\gamma) = E[g(\epsilon+\gamma)]$。$\gamma$ 的增加，总是会降低双方获胜的概率，并且其可能性降低的值正好是 $1/\theta$ 乘以代理人努力工作所增加的获胜机会，

$$\partial P/\partial\gamma = -\bar{g}(\gamma) \tag{33}$$

定理6：如果扰动项的分布是对称的，或者在更加一般的情况下，如果 $g(\bar{\epsilon}) \geqslant g(\underline{\epsilon})$，其中 $\bar{\epsilon}$ 和 $\underline{\epsilon}$ 是 ϵ 的最大值和最小值，那么引入严格正的领先距离能够改进锦标赛制的结果。

证明：证明过程首先要解出 γ 的值，然后再证明在定理6的条件下，最优的值是正的。首先，在对称均衡里有，

$$\partial\bar{g}(\gamma)/\partial\gamma = E[g'(\epsilon + \gamma)] \tag{34}$$

γ 的值的变化会影响到代理人的努力付出以及他的期望产出水平，

$$\partial\bar{Y}/\partial\gamma = -\bar{g}'\Delta UE[\theta^2/V'']/\{1 - \bar{g}\Delta U'E[\theta^2/V'']\} \tag{35}$$

而这种变化的符号由 $\bar{g}'$ 的符号决定。当奖金额 x 被最优的确定后，γ 的值的增加只能通过他对努力程度，期望产出和获胜机会的效应来施加影响

$$\begin{aligned}\partial W/\partial\gamma = &-\bar{g}(\gamma)[U(\bar{Y} + x) + U(\bar{Y} - x) - 2U(\bar{Y})]\\ &+ [PS + (1 - 2P)U'(\bar{Y}) - \bar{g}\Delta U]\bar{Y}'\end{aligned} \tag{36}$$

因为 $U(Y)$ 的凸性，所以第一项的值为正。当 $\gamma=0$，$P=1/2$。第二项的符号同 $\bar{Y}'$ 是一样的，因为由式（24）可得 $(1/2)S > \bar{g}\Delta U$。在定理6的条件下，$\bar{Y}'$ 同样为正值

$$\partial\bar{g}(0)/\partial\gamma = 1/2[g^2(\bar{\epsilon}) - g^2(\underline{\epsilon})] > 0 \rightarrow \bar{Y}' > 0 \tag{37}$$

因此，当 $\gamma=0$，$\partial W/\partial\gamma>0$，从而引入领先距离是值得的。更一般地说，要解出最优的 γ 的值必须解出由式（36）$\partial W/\partial\gamma=0$ 的解。

在一个企业内施行竞争性报酬体系时，引入领先距离还可以提高竞赛的可持续性。获胜的工人被提升到了管理层的高端，他们就有能力重新设计激励安排。通常这些高层管理者都拥有好于常人的运气。他们中的许多人都明白如果竞赛评估过程是反复进行的话，他们是有可能失去他们的职位的。因此，如果想要竞赛能够持续的话，就必须使得技术和努力程度在取胜的过程

中的影响力能够超过运气因素。纳尔巴夫（1982）就在有领先距离条件下的非对称均衡中找到了一个解决办法。如果有的代理人更加努力工作，那么多数情况下获胜者都是更加努力的人，只有很少的情况下是由运气决定的。

有许多竞争者的竞赛。增加参赛者的人数有两种效果：其一是由产出传递的信息量增加了（就如同在第四部分见到的充分统计量定理的结果）；其二是薪酬安排的决策范围更广泛了。我们将从一个简单的一般情况开始，胜者只有一种奖金；而把有 n 种奖金的锦标赛制推迟到第十二部分再讨论。当存在 n 个人为了一个奖金而竞争时，可以认为每个人都投入了 x 的努力，因而胜者的奖金就是 nx。在对称均衡的情况下，每个人的期望效用是

$$W = U(\bar{Y} - x) + (1/n)\Delta U(n) - E[V(\mu)] \tag{38}$$

其中 $\Delta U(n) = U(\bar{Y} - x + nx) - U(\bar{Y} - x)$

同前面一样，我们可以计算出努力水平的增加的期望收益。假设其他每个人都付出努力水平为 $\bar{\mu}$，而第 i 个代理人付出的努力水平为 μ_i；这样，如果他观察到的真实值，那么只有当对于所有的 j，都有

$$\epsilon_j \leqslant \epsilon_i + \theta(\mu_i - \bar{\mu}) \tag{39}$$

的时候，他才可能获胜。而这种情况发生的概率为 $G(\epsilon_i + \theta(\mu_i - \bar{\mu}))^{n-1}$。获胜的概率就是此概率条件于 ϵ_i 期望，

$$P = \int G(\epsilon_i + \theta(\mu_i - \bar{\mu}))^{n-1} g(\epsilon_i) d\epsilon_i \tag{40}$$

在对称均衡处，$P = 1/n$，而努力工作的边际获胜机会的增加 $\partial P/\partial u$ 等于

$$\partial P/\partial \mu = \theta \bar{g}(n) = \theta(n-1)\int [G(\epsilon)]^{n-2} g^2(\epsilon) d\epsilon \tag{41}$$

委托人能够决定最优的奖金水平 $x^*(n)$ 使得每一个代理人都付出与最优情况下相同的努力水平。只要令

$$\bar{g}(n)\Delta U(n) = U'(\bar{Y}) \tag{42}$$

像式（42）一样重写式（19），注意只要 $\bar{g}(n)$ 趋近于零的速度比（$1/n$）稍慢，那么在极限时 $\Delta U(n)/n$ 一定也为零。

$$\lim_{n\to\infty} n\bar{g}(n) \cdot (\Delta U(n)/n) = U'(\bar{Y}) \Rightarrow \left[\lim_{n\to\infty} n\bar{g}(n) = \infty \to \lim_{n\to\infty}(\Delta U/n) = 0\right] \tag{43}$$

回忆式（11）中的期望效用，如果效用函数形式是线性渐进的（如此边际效用就有下界）[7] 而且 $\Delta U(n)/n$ 也趋向于零，那么奖金额 x 也趋向于零而

---

7 否则，如果 $U'(Y)$ 趋向于零，那么对于任何有限的 x 值，$\Delta U(n)/n$ 也趋向于零，这样还是存在着非凸性问题。

且期望效用的值趋向于最优值：

$$\lim_{n\to\infty} \Delta U(n)/n > nxU'(\bar{Y}+nx)/n = xU'(\bar{Y}+nx) = 0 \to \lim_{n\to\infty} x^*(n) = 0 \quad (44)$$

而

$$\lim_{n\to\infty} W = U(\bar{Y}) - E[V(\mu^*)] \quad (45)$$

就是最优效用函数值。

但是，非凸性问题会破坏对称均衡。代理人如果选择不工作只会以 $1/n$ 的概率损失 ΔU，而节省了 $V(0) - V(\mu^*)$ 的效用损失。代理人如果选择不工作，他的期望效用为

$$\lim_{n\to\infty} W_{|\mu_i=0} = U(\bar{Y}) - V(0) > U(\bar{Y}) - E[V(\mu^*(\theta))] \quad (46)$$

乍看起来，增加锦标赛的规模是很有吸引力的，因为它可以用更小的风险提供更好的激励；然而不幸的是，增加锦标赛的规模同样会改变奖金额使得破坏均衡变得更加有吸引力。

惩罚。莫里斯（Mirrlees，1974）在讨论如何惩罚不能完成任务的工人时，提出了一个关于非凸性问题的解。在他的框架下，这意味着对于最落后的工人要施以惩罚。随着代理人的数目越来越多，成为最后一名的机会变得非常之小。但是决定惩罚水平的关键是边际条件。只要边际激励不是下降得太快，委托人就无须对失败者施加越来越重的惩罚。

从数学上讲，倘若交换 x 和 $-x$ 以及获胜的概率和失败的概率，那么惩罚模型就应该与奖金模型一致。在对称均衡处，有

$$W(n) = U(\bar{Y}+x) - \Delta U(n)/n - E[V(\mu(\theta))] \quad (47)$$

其中 $\Delta U(n) = U(\bar{Y}+x) - U(\bar{Y}+x-nx)$。

同前面一样，当只存在单一奖金的时候，最优条件决定的产出为

$$-\theta\Delta U(n)\tilde{g}(n) - V'(\mu(\theta)) = 0 \quad (48)$$

其中 $\theta\tilde{g}(n)$ 是努力工作后失败机会降低的边际值

$$\tilde{g}(n) = -(n-1)\int[1-G(\epsilon)]^{n-2} g^2(\epsilon)d\epsilon \quad (49)$$

一般而言，当误差项分布在 $[-\infty, \infty]$ 的支持上时，随着参赛者人数变得很大，失败机会的边际改变值 $\theta\tilde{g}$ 也会趋近于零。不过只要其收敛速度小于 $1/n$（即当 $n\to\infty$ 的极限时有 $n\tilde{g}(n) = -\infty$），代理人的期望效用就将趋近于最优解的水平，而且存在一个纯策略纳什均衡。

为了复制最优解的努力水平，选择 x 满足

$$U'(\bar{Y}) = -\tilde{g}(n)\Delta U = -n\tilde{g}(n)[\Delta U(n)/n] \quad (50)$$

当 $-n\tilde{g}$ 趋近于无穷，奖金额一定是使得 $\Delta U(n)/n$ 趋近于零。如果 $\Delta U(n)/n$ 趋近于零，那么奖金额也一定趋近于零

$$\Delta U(n)/n > nxU'(\bar{Y}+x)/n = xU'(\bar{Y}+x) \quad (51)$$

因此，福利水平达到了最优解的水平，

$$W = \lim_{n\to\infty} U(\bar{Y}+x) - \Delta U(n)/n - E[V(\mu^*(\theta))] = U(\bar{Y}) - E[V(\mu^*(\theta))] \quad (52)$$

在惩罚的框架下，所有代理人都付出 $\mu^*(\theta)$ 的努力水平构成了纳什均衡解。在只有一个胜利者的模型里，偷懒的成本最多也不过是 $\Delta U(n)/n$，而他是随着 n 的逐渐变大而趋近于零的。这里偷懒有可能造成大到 $(n-1)\Delta U(n)/n$ 的损失，后者随着 n 的增大最多可以达到 $\Delta U(n)$ 偷懒不大可能毁掉参赛者所有的获胜机会。代理人的最大化问题是凸性的，这是因为：（i）偷懒的人越多，代理人节省的效用损失就越少，因为依据假设他的边际效用损失是随着他的努力水平而递增的；（ii）代理人越是偷懒，他失败的边际的 ϵ 值也越小，因而他获胜的机会更大了，他失利机会增加的边际成本是上升的。更加正式地说，对于 $\text{sign}(g'(\epsilon)) = -\text{sign}(\epsilon)$ 的单峰分布而言，代理人越是偷懒，他失败的概率降低的边际量的负值越小。

在胜利者只有一种奖金的竞赛之中，代理人越是偷懒，他获胜的边际的 ϵ 值越大，从而他获胜的机会也越小，因此作弊的边际收益降低，于是产生了非凸性问题。图 1 阐释了这个问题。然而企图将莫里斯提出的把偶尔并且严肃的惩罚转换为高额而不太可能的奖金是一种冒险的想法。霍姆斯特姆（Holmstrom，1982，Theorem 4；1977，p. 242）提出“如果代理人的效用函数是线性渐进的形式，而且委托人是风险中性的，那么完全可以照搬前面的命题来推出结论说，通过高额但是不太可能的奖金是可能引致出接近最优的努力水平的”。这个问题也可能是满足局部凸性的（从而二阶条件是局部厌足的）。然而，如果要接近最优的努力水平，期望的奖金额就必须为零。代理人如果作弊最多也就是损失了趋近于零的期望奖金额，但是却节省了所有努力所产生的负效用。因此，最优解不是全局最优的。相比之下，在存在着惩罚条件下，代理人越是作弊，他的成本也越高。如果他作弊足够多，那么他的期望惩罚值就将从零上升为一个确定的值 $-\Delta U(n)$。

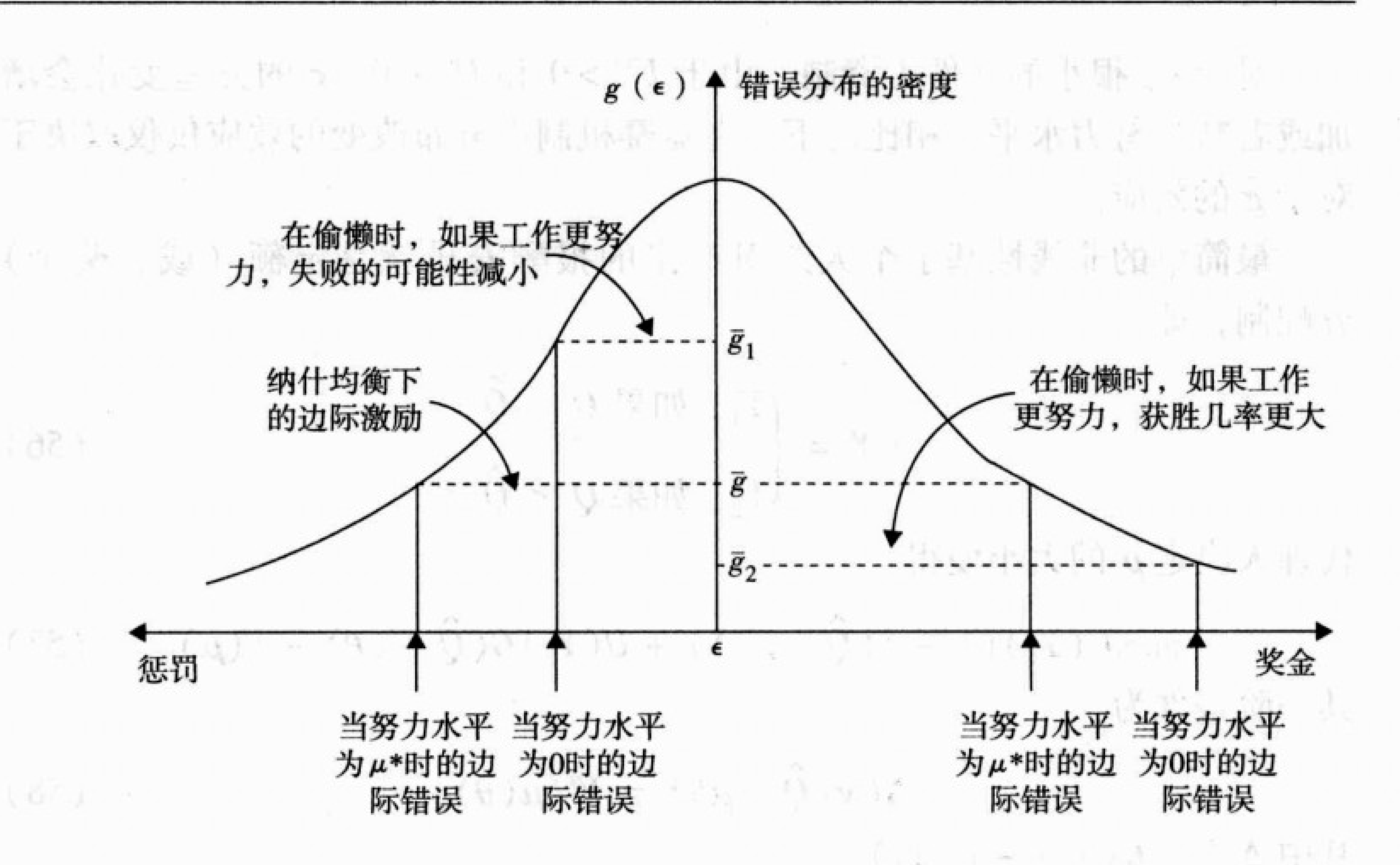


图 1　当惩罚存在时，欺骗使得边际 ϵ 变为 0，到达一个密度较高的区域。当奖金存在时，欺骗使得边际 ϵ 向上移动，远离原点，到达一个密度较低的区域。

九、基于个人绝对产出的安排：计件工资率和定额分配制

最简单而又最常用来研究的两种薪酬安排就是线性计件工资率和定额分配制。这里我们把这些报酬机制与竞赛机制作个比较。

零利润条件意味着代理人在线性计件工资方案下的薪酬是

$$Y_i = \alpha Q_i + (1-\alpha)\overline{Q} \tag{53}$$

个人得到的报酬是由期望产出加上一个基于其自身产出与集体产出之差的激励性奖金。由一阶条件可以决定努力水平，[8]

$$\alpha\theta E[U'(\alpha Q_i + (1-\alpha)\overline{Q})] = V'(\mu(\theta)) \tag{54}$$

而在竞赛条件下，μ 能够对 θ 的值自动调整。但是其值要小于在竞赛条件下或者最优均衡条件下的值

$$\frac{d\mu}{d\theta} = \alpha[E(U' + \alpha\theta\mu U'')]/[V'' - \alpha^2\theta^2 EU''] \tag{55}$$

μ 的变化方向甚至可能与 θ 相反。当 σ_θ^2 和 σ_ϵ^2 很小的时候，那么对于对数效用函数有 $d\mu/d\theta \approx 0$。

8　参见 Arnott and Stiglitz（1980）的论述，可能会存在不满足二阶条件的问题。

对于 σ_ϵ^2 很小的计件工资制，由于 $U''' > 0$ 和 $U''' < 0$。ϵ 的方差变化会增加或者减少努力水平。相比之下，在竞赛机制中分布改变的效应仅仅取决于对于 $\bar{g}$ 的效应。

最简单的非线性基于个人绝对产出的报酬安排就是定额（或者奖金）分配制，即

$$Y = \begin{cases} Y_1, \text{如果 } Q > \hat{Q} \\ Y_2, \text{如果 } Q < \hat{Q} \end{cases} \tag{56}$$

代理人决定 μ 的大小使得

$$\max U(Y_1)(1 - G(\hat{Q} - \mu\theta)) + U(Y_2)G(\hat{Q} - \mu\theta) - V(\mu) \tag{57}$$

其一阶条件为

$$\theta\Delta\tilde{U}g(\hat{Q} - \mu\theta) = V'(\mu(\theta)) \tag{58}$$

其中 $\Delta\tilde{U} = U(Y_1) - U(Y_2)$。

采用这两种如此简单的定额分配和奖金机制立刻会带来困难。首先，即使个人是风险中性的，奖金安排也无法引致一级最优解的努力水平（因为 $g' \neq 0$ 除非 g 是均匀分布）；其次当 θ 很小时，存在着非凸性问题。当分配定额很高而任务很难完成时，代理人可能会觉得竞逐奖金是不值得的。

十、竞赛制与计件工资制的比较

当产出具有高度波动性的时候，竞赛制很可能是优于计件工资制的。代理人会截去极端的可能性，把风险的大小限制于 $\pm x$ 之内。在计件工资制里，工人必须接受一个小概率的低收入可能性。

而当代理人是风险中性的情况下，竞赛制与计件工资制都能够实现最优解的努力水平。在计件工资制里，委托人制定了计件工资率 α 等于 1。此时不存在风险的福利损失，激励水平也同最优解的情况下相同。而在竞赛制中，委托人通过对每一个自然状态下都按照式（19），$x^* = 1/(2\bar{g})$ 来设定奖金额就能够实现最优解的努力水平。在激励出 $\mu^*(\theta)$ 的同时，奖金不会影响到代理人的效用水平。这是因为期望为零而且代理人是风险中性的。

尽管有很多种方法引入风险规避，我们仅仅考虑二次型的效用函数的简单例子，于是有

$$W = E\{a + bY - c(Y - Y^*)^2 - \mu^2/2\} \tag{59}$$

其中 Y^* 是最优解的期望产出。当然 $c=0$ 对应于风险中性。我们首先考虑当 c 由零增加对于期望效用的影响。

$$dW^C/dc_{|c=0} = (dW^C/dx)(dx/dc) + (\partial W^C/\partial \mu)(\partial \mu/\partial c) + \partial W^C/\partial c \tag{60}$$

因为奖金额是最优的，所以第一项的值为零，努力水平的决定依赖于下式：

$$\mu = \theta \bar{g} \Delta U = \theta \bar{g} 2bx \tag{61}$$

当 $c=0$ 时最优的努力水平不受 c 的变化的直接影响，于是第二项的值也为零。风险规避导致的效用损失就和风险的变化 x 成比例。此结果并不依赖于 σ_θ^2 或者 σ_ϵ^2 要很小的条件。当 $c=0$ 时 $x=1/[2\bar{g}]$ 并且，

$$dW^C/dc_{|c=0} = \partial W^C/\partial c = -E[(Y-Y^*)^2] = -x^2 = -1/[4\bar{g}^2] \tag{62}$$

在计件工资制里，

$$dW^p/dc_{|c=0} = (dW^p/d\alpha)(d\alpha/dc) + (\partial W^p/\partial \mu)(\partial \mu/\partial c) + (\partial W^p/\partial c) \tag{63}$$

当 α 取得最优值 1，α 值的微小改变不会有影响。由于工人选择了最优的努力水平，不努力工作导致的消费水平降低正好被休闲的增加抵消了，$\partial W^p/\partial \mu=0$。同样地，只有 c 的直接效应才是重要的。产出的方差等于 ϵ 的方差加上 $\theta\mu$（θ）的方差。当 $c=0$ 时 $\mu=b\theta$，这样 $\theta\mu$（θ）的方差就是 $b^2\mathrm{Var}$（θ^2）：

$$dW^p/dc_{|c=0} = (\partial W^p/\partial c) = -(\sigma_\epsilon^2 + b^2\mathrm{Var}(\theta^2)) \tag{64}$$

在 c 很小的情况下可以很容易地比较竞赛制与线性计件工资率的福利水平，如果

$$(\sigma_\epsilon^2 + b^2\mathrm{Var}(\theta^2)) > 1/[4\bar{g}^2]\text{，那么 } W^C > W^p \tag{65}$$

显然当 Var（θ^2）的值足够大，竞赛制永远要优于线性计件工资率。即使当 Var（θ^2）$=0$，也可能找到 ϵ 的某种分布使得竞赛制更优（例如在 $[-1, 1]$ 上为 g（ϵ）$=5/2\epsilon^4$，其他点处为零的分布）。[9]

十一、相对绩效

竞赛的特别之处在于决定薪酬的因素只有名次。这里考虑几个简单的薪酬安排，其中相对绩效的大小进入了薪酬安排。这就使得我们能够获得关于

9　由充分统计量定理我们知道如果 θ 是常量的话，最优的非线性计件工资率要优于竞赛制。事实上如果 θ 是常量的话，的确有可能证明某个奖金机制占优于线性计件工资率，而当效用函数为二次型并且代理人显示出轻微的风险规避倾向时，线性计件工资率要远远劣于竞赛制。

何时竞赛制会更优的一些直觉 。

考虑这样的一个薪酬安排

$$Y_i = \beta(Q_i - Q_j) + \bar{Y} \tag{66}$$

代理人获得的奖金额是个定值 $\bar{Y}$ 加上一个关于两个代理人的产出之差的线性函数。在对称均衡处有

$$Y_i = \beta(\epsilon_i - \epsilon_j) + \bar{Y} \tag{67}$$

于是个人的收入方差正好是 $2\beta^2\sigma_\epsilon^2$。

两个代理人产出与 θ 的相关性是一致的，但是与 ϵ_i 是相互独立的。相对于计件工资率，将另一个代理人的产出作为负相关的“资产”消除了所有由 $\theta\mu$（θ）的变化产生的收入不确定性，而代价仅仅是附加了另一个 ϵ 的方差。如果 E（ϵ^2）比 E［（$\theta\mu$（θ））^{2}］足够小的话，那么此种安排就是更优的。

薪酬安排在霍姆斯特姆（1982）的充分统计量结论里有深入的阐述。当代理人的数目非常大的时候，可以将报酬挂钩于个人产出与平均产出的差值。当平均产出收敛于其真实值 $\theta\mu$（θ）的时候，代理人仅仅需要承担由他们自己的个人扰动 ϵ_i 所带来的风险。

通过按照下式来选择 β 值，基于两个代理人的产出差距大小的薪酬安排就可以实现最优的努力水平，

$$\beta E[U'(\beta(\epsilon_1 - \epsilon_2) + \bar{Y})] = U'(\bar{Y}) \tag{68}$$

然而，倘若我们将薪酬安排与个人产出与平均产出的比率挂钩的话，如果代理人的数目很大，那么结论就不是如此了，$Y_i = f$（$Q_i/\bar{Q}$）。

$$\frac{Q_i}{\bar{Q}} = \frac{\mu_i}{\mu^*} + \frac{\epsilon_i}{\theta\mu^*} \tag{69}$$

个人选择 μ_i 值使得，

$$\frac{1}{\mu^*}E[U'f'] = V'(\mu_i) \tag{70}$$

这里 μ 仅仅通过 f 的效应随着 θ 而变化，因此，

$$\frac{d\mu}{d\theta} = \frac{-E\{(U''f'^2 + U'f'')\epsilon_i\}/\mu^2\theta^2}{V'' - E(U''f'^2 + U'f'')/\mu^2} \tag{71}$$

如果 $\sigma_\epsilon^2 = 0$，相比最优解的情况 μ 是不随 θ 而变化的，前者的情况下个人的生产能力（θ）上升会导致努力水平的增加。这种相对绩效的评价标准，缺乏我们前面所说的“灵活性”。

十二、扩展：多重奖金的锦标赛制

当锦标赛的参赛者人数非常多时，利用几种不同的奖金额所增加的灵活性就可以改进单一赢家或者单一输家的锦标赛制。令代理人 i 以第 j 名结束比赛的概率为 $P(Q_i=j)$，则

$$P(Q_i = j) = \int \frac{(n-1)!}{(n-j)!(j-1)!} g(\epsilon) G\{\theta(\mu_i - \bar{\mu}) + \epsilon\}^{j-1} [1 - G\{\theta(\mu_i - \bar{\mu}) + \epsilon\}]^{n-j} d\epsilon \tag{72}$$

其中 $\bar{\mu}$ 是其他 $n-1$ 个代理人的努力水平。在对称均衡之中（$\mu_i=\bar{\mu}$），代理人 i 最后得到任何名次的可能性都是相等的 $P(Q_i=j)=1/n$，当获得第 j 名的奖金为 $\tilde{\omega}_j$，那么期望效用就是

$$W = (1/n)\left[\sum_{j=1}^{n} U(\tilde{\omega}_j)\right] - E[V(\mu(\theta))] \tag{73}$$

努力水平由下式决定

$$\sum_{j=1}^{n} \frac{\partial P(Q_i = j)}{\partial \mu_i}[U(\tilde{\omega}_j)] - V'(\mu(\theta)) = 0 \tag{74}$$

在对称均衡条件下努力工作会增加（或减少）获得第 j 名的机会，

$$\partial P(Q_i = j)/\partial \mu_i = \theta \int \frac{(n-1)!}{(n-j)!(j-1)!} g^2(\epsilon)(1 - G(\epsilon))^{n-j-1} G(\epsilon)^{j-2} \times \{(j-1)(1 - G(\epsilon)) - (n-j)G(\epsilon)\} d\epsilon \tag{75}$$

很容易想像，如果有足够多的不同奖金额度的话，锦标赛制就能够达到任何基于个人绝对产出的非线性薪酬安排的努力水平。方法就是将整个非线性奖金安排化为百分比形式，然后通过在不同百分比之间平摊的报酬结构来逐步接近（如同格林和斯托基（1981）的方式）。竞赛中获得第 j 名的代理人的奖金就等价于位于百分之 $100j/n$th 的代理人的收入。尽管在对称均衡处这两种安排达到的期望效用水平是相同的，但是他们的边际激励水平并不一样。因此，一般而言在纳什均衡里等价性并不是能够持续的。在竞赛机制中，由方程式（74）和式（75），最优的努力水平满足

$$V'(\mu(\theta))/\theta = \text{常数（与 } \theta \text{ 不相关）} \tag{76}$$

对于一般的非线性计件工资率安排 $Y(Q_i)$，决定最优努力水平的一阶条件对努力水平的效用损失并不和 θ 成比例，

$$\theta E[U'\{Y(\theta\mu(\theta) + \epsilon)\}Y'(Q)] = V'(\mu(\theta)) \tag{77}$$

非线性计件工资率安排只能够在下面条件下复制竞赛的效果（反之亦然）：

(i) 对于 θ 不是变量（格林和斯托基（1981）考虑的模型中一般扰动不会影响边际努力的产出）或者（ii）引致代理人使得他表现出风险中性的特征。$Y(Q)$ 由 $U(Y(Q)) = z_1 + z_2 Q$ 决定。

在 n 种奖金条件下的名次锦标赛里有两种特别的情况是可以相对容易地求出最优的奖金额的。

(1) 均匀分布，$g(\epsilon) = \bar{g}$ $\{\epsilon: [-A, A]\}$

(2) 指数分布，$g(\epsilon) = e^{\epsilon - 1}$ $\{\epsilon: [-\infty, 1]\}$。

当扰动服从均匀分布时，奖金 2，…，$n-1$ 不会对工作激励产生影响。而得到内点解或者排除内点解的可能性是相等的。因此，努力工作以增加获得任何内点解奖金额的可能性为零。所有的工作激励都必然是由 $\tilde{\omega}_1$ 和 $\tilde{\omega}_n$ 产生的。为了最小化风险，委托人将所有的内点解奖金额都设定为相等是最优的 $\tilde{\omega}_2 = \tilde{\omega}_3 = \cdots = \tilde{\omega}_{n-1}$。即使有 n 种可能的奖金额，在扰动服从均匀分布条件下，最一般的报酬安排也只需要三种不同的奖金额：一个末等奖，一个一等奖，再加上一个不变的中等奖。只有在当名次变化很大时，委托人不得不采取更多数量的惩罚来确保所有代理人都选择了局部最优解的情况下能够得到验证，此结果才是正确的。

当扰动服从指数分布时，可以证明提高努力水平会减少除了一等奖之外获得任何其他奖项的可能性。因为其他 $n-1$ 种奖金的边际激励水平是完全一样的，这些奖金必须等值以确保风险最小。于是采用竞赛制的委托人只需要设定另外两种不同的奖金额，就能够达到次优解的努力水平。

尽管和前面一样，这里还需要验证非凸性问题，但是这些例子已经显示出研究简单的报酬结构，可以用来很好地估计报酬结构更加复杂的名次锦标赛制。

十三、扩展：当产出不是努力水平的线性函数

截至目前的分析，仅仅考虑了产量 Q_i 是努力水平的线性函数的情况 $Q_i = \theta\mu_i + \epsilon_i$，更加一般的结构应该考虑产量是努力水平的函数的情况。令 ϵ 为测量误差，那么，

$$(\text{i})\ Q_i = F(\theta\mu_i) + \epsilon_i$$

第二种情况（也许这个情况更加有意思）是有理由将 ϵ 视为一个运气因素

$$(\text{ii})\ Q_i = F(\theta\mu_i + \epsilon_i)$$

一旦公共的环境因素 θ 已知，那么努力水平也就决定了。

在第一种情况下，最优解的特征是

$$\bar{Y} = E[F(\theta\mu_i(\theta)] \tag{78}$$

以及对于给定μ_i（θ），

$$\theta U'(\bar{Y})F'(\theta\mu_i(\theta)) = V'(\mu_i) \tag{79}$$

在锦标赛制里提高努力水平可以增加获胜的机会。第一个代理人获胜的条件是$F(\mu_1\theta) - F(\mu_2\theta) + \epsilon_1 > \epsilon_2$。而在对称均衡解处$\mu_1 = \mu_2$，努力工作提高获胜的概率，其比例为$F'$的一个比例

$$\partial P(\mu,\mu,\theta)/\partial\mu_i = P_{\mu_i} = \theta F'(\theta\mu_i(\theta))\int g^2(\epsilon)d\epsilon = \theta F'(\theta\mu_i(\theta))\bar{g} \tag{80}$$

以及给定μ_i（θ），有

$$P_{\mu_i}\Delta U = V'(\mu_i(\theta)) \tag{81}$$

委托人通过设定奖金额满足$\bar{g}\Delta U = U'(\bar{Y})$依然可以实现最优的努力水平。其分析过程与前面类似。

考察ϵ是一个运气因素的含义会更有趣，在此情况下，最优解满足下列条件：在给定μ_i[10]情况下，有

$$\theta U'(\bar{Y})\bar{F}'(\theta\mu_i(\theta)) = V'(\mu_i(\theta)) \tag{82}$$

在此情况下的锦标赛制里，生产函数F与竞赛优胜者的决定过程是不相关的，因为它不过是$\mu\theta + \epsilon$的一个单调变换而已。即，

$$F(\theta\mu_1(\theta) + \epsilon_1) > F(\theta\mu_2(\theta) + \epsilon_2) \Leftrightarrow \theta\mu_1(\theta) + \epsilon_1 > \theta\mu_2(\theta) + \epsilon_2 \tag{83}$$

因此，决定努力水平的一阶条件是：对于给定μ_i（θ）

$$\theta\bar{g}\Delta U = V'(\mu_i(\theta)) \tag{84}$$

不幸的是，这个表达式中没有F'项，因此除非F是线性函数，否则锦标赛制就无法达到最优的努力水平。所以即使代理人是风险中性的，也不能保证达到最优的努力水平。

十四、奖金影响技术的选择

在竞赛机制条件下研究技术的选择问题，最初是由斯蒂格利茨和魏斯（Weiss，1981）在研究信贷配给问题时提出的。银行的贷款就好像是定额分配与计件工资的一个组合。如果代理人挣得的钱不足以偿付利息，那么他们

10　其他的变量由$\bar{F} = E[F(\theta\mu(\theta) + \epsilon)]$定义，预期函数是对$\epsilon$求预期，对于预期函数$\bar{Y} = E[\bar{F}]$是对$\theta$求预期。

就会破产获得零收入。如果代理人的产出很高，那么他会保留超出利息支付的部分。利率（分配定额的水平）的升高会导致借款人采取更加冒险的策略（Rothschild and Stiglitz，1970）。

信贷配给问题可能导致借款银行不愿意以高利率借款给高风险的客户。但是在其他的情况下高的风险也许是受欢迎的。与信贷市场相比，竞争性薪酬安排（即竞赛制或者定额分配制）相对于计件工资率的一个潜在优势是：它鼓励企业家抛弃原有的厌恶风险的倾向，转而选择高风险高回报的项目。这种优势的作用在研发领域是尤其巨大的，在这些领域里奖金（以专利的形式）可以大大缩短开发时间。

我们所考察的薪酬安排都缺乏足够的工具变量来同时控制投入（努力水平）和技术。

十五、结论

社会也许会发现竞争的结果——失败者的产生——是不能接受的。基于个人绝对产出的薪酬安排的一大优势就在于它能够保证每个人都有高的产出，达到分配的定额，并且获得报酬。相比之下，当委托人不知道如何设定分配定额，从而将报酬与相对绩效挂钩的时候，总会有人成为最后一名，而沦为失败者。惩罚所带来的问题在于它能够自我维持下去。失败者的士气受挫，无以为续，于是以后只能更加失败。但是仅仅靠奖金是不足以调动努力的积极性的，因为对于很大的一个群体而言，期望的奖金额太小，不足以制止怠工。

竞争机制在所有参赛者都是同质的情况下运转最佳。实施惩罚的一个困难之处在于失败者往往不仅仅是运气不好，而且常常是能力也不够。他们在竞争的时候，往往比普通的工人付出了更多的努力。更糟的是，如果有人一定会输的话，那么其他的人也就失去了努力工作的动力了。一旦知道了各人的相对能力，障碍赛制（比如高尔夫巡回赛）就能够恢复公平竞争的比赛环境。

在非竞争性的生产安排条件下，相当大的激励是依靠社会因素而不是经济因素实现的。代理人不愿意偷懒是因为偷懒不被社会接受。由于每个工人都认为：如果只能获得失败者的收入，那么还不如不努力工作，所以在奖金不高而且代理人数很多的时候，依然会产生非凸性问题。尽管假设经理（委托人）无法监督代理人的努力水平，但是工人们却可以相互监督。工人们也许不情愿指出干的略少的同伴，但是偷懒会导致降低努力水平的不连续

性。如果可以惩罚偷懒的个人而不是给他失败者的报酬，那么即使奖金不高也同样可以存在纯策略纳什均衡解。由于存在着获得一级最优水平的可能性使得每个人都有越来越大的激励不接受偷懒的个人（Akerlof and Soskice，1976，*Theory of Sanctions*）。

竞争性薪酬安排具有很明显的优势，然而似乎却未被广泛采用。这可能是由于工人的满意程度造成的，而后者往往被传统的经济模型所忽略。这一类想法（即工作环境、成员的同质性程度）在分析企业之间的竞争时，往往不如在分析企业内部竞争时那么重要。

在充满竞争的环境里，比如专利竞赛，合作可能会产生技术性的回报，例如共享信息。在完全竞争的市场中是不存在合作激励的。相反，破坏别人的成果甚至是有利可图的，只要不会伤及自身。计件工资率会鼓励工人们在有共同利益的时候相互合作，而这一点可能是非常重要的。当委托人的报酬函数是对代理人的能力水平的凸函数时，采取竞赛制会产生进一步的决策权衡。尽管代理人的工作激励是增加了，但是委托人却因为几个项目同时运行存在着效率损失（Nalebuff and Varian，1981）。

本文中模型提出的竞争是具有不同作用的——真正的竞争，在通常使用的语义下，而不是在许多新古典经济学已经开始使用的那个静态的特殊含义下——在存在着关于不同工作的难易程度的不完美信息条件下，直接观察投入的成本异常大，同时也很难准确地测量出产出水平。采用比较体系的优势在于对于环境的变化，它们比其他的薪酬安排拥有更大的灵活性以及更大的适用性。

附录　混合策略均衡

我们已经在正文中证明了在单一获胜奖金的竞赛制中，不论当参赛者人数变大还是当由 ϵ 造成的方差（风险）趋近于零，都不存在纯策略纳什均衡。在两种情况下期望奖金的值都太低，使得竞逐奖金是不值得的。因此两个代理人都会选择不工作而挣得败者报酬。当 ϵ 的支持是有界的时候，将绩效评价标准同竞赛制结合起来可以部分地解决此问题（Stiglitz，1975）。

假设在极端的情况下 $\sigma_\epsilon^2=0$。显然，任何代理人都可以通过将努力水平提高到比对手稍微高一点的水平上来确保胜利。那么每一个代理人的反应函数都是不连续的。因此，不存在纯策略纳什均衡。令 $H(\mu)$ 为面对竞争者最大努力水平的概率函数，付出努力水平为 μ_i 的个人的期望效用一定是常数

$$U(\overline{Y}-x)+H(\mu_i)\Delta U-V(\mu_i)=k \tag{A1}$$

显然，如果

$$k=U(\overline{Y}-x)-V(0) \tag{A2}$$

而且

$$H(\mu)=[V(\mu)-V(0)]/\Delta U \tag{A3}$$

个人对于努力水平是无差异的。令个人的能力水平的分布函数为 $R(\mu)$，对于 n 个代理人，面对 $n-1$ 个对手，

$$H(\mu)=R(\mu)^{n-1} \tag{A4}$$

$$R(\mu)=\{[V(\mu)-V(0)]/\Delta U\}^{1/[n-1]} \tag{A4a}$$

期望产出为

$$\overline{Y}=\int_0^{\tilde{\mu}}\mu r(\mu)d\mu=\int_0^{\tilde{\mu}}[1-R(\mu)]d\mu \tag{A5}$$

其中最大努力水平 $\tilde{\mu}$ 满足

$$V(\tilde{\mu})-V(0)=\Delta U \tag{A6}$$

最优可行的混合策略均衡由奖金的选择 x^* 决定，满足最大化期望效用 k。由式（A2）可以得出

$$U'(Y-x)[\overline{Y}'-1]=0 \tag{A2a}$$

因此，我们需要寻找满足 $\overline{Y}'=1$ 的解。由式（A5）和式（A6），得

$$d\overline{Y}/dx=[(\tilde{\mu}-\overline{Y})S]/[(n-1)\Delta U-(\tilde{\mu}-\overline{Y})\Delta U'] \tag{A7}$$

为了给出一个完整的混合策略均衡解的例子，进一步假设

(i) $V(\mu)=\mu^2/2$ 并且

(ii) $U(x)=x$ 风险中性

之所以选择风险中性的效用函数是因为分析起来简单，而且相比采取计件工资率所达到的一级最优结果，一般较少的混合策略均衡解此时可以很明显地找到。将其代入式（A5）和式（A6），得

$$\tilde{\mu}^2/2=2x \tag{A6a}$$

$$\overline{Y}=\int_0^{\tilde{\mu}}\{1-(\mu/\tilde{\mu})^{2/(n-1)}\}d\mu=2\tilde{\mu}/(n+1) \tag{A5a}$$

在点 x^*，$\overline{dY/dx}=1$ 由式（A7）得到

$$(n-1)x=[\tilde{\mu}-\overline{Y}] \tag{A7a}$$

因此，在最优混合策略均衡处，

$\tilde{\mu}=4/(n+1)$，$x^*=4/[(n+1)^2]$，而且 $\overline{Y}=8/[(n+1)^2]$，

期望效用为 $EU=U(\overline{Y}-x^*)-V(0)=\overline{Y}-x^*=4/[(n+1)^2]$

(A8)

在最优解处，工人们选择μ以便让下式最大化

$$EU = \mu - \mu^2/2 \to \mu^* = 1 \text{ 以及 } EU = .5 \quad (A9)$$

我们观察到在只有两个工人的情况下，混合策略均衡可以相当地靠近点$\left(\frac{4}{9}\Big/\frac{1}{2}\right)$最优的效用水平。然而随着参赛者的人数越来越多，这种竞争形式就变得越来越残酷，直到所有的消费者剩余都被耗尽了。由于工人们知道一旦卷入到这场竞争的角逐之中，所有的剩余都会被夺走，如此一来他们就不会签署这么一个合同（如果还存在着其他有正的消费者剩余的合同的话）。当σ_ϵ^2变得很小，竞赛的条款可以这样设定使得残酷的竞争只有在某些θ值的条件下才会发生，而对于其他的（较小的）θ值，会发生传统的竞争。包含计件工资率的合同一样可以消除这样的两难窘境（更加详细的关于专利竞赛条件下的讨论请参见 Gilbert and Stiglitz，1979）。

参考文献

AKERLOF，G，“The Economics of Caste and of the Rat Race and Other Woeful Tales，” *Quarterly Journal of Economics*，Vol. 90（1976），pp. 599 –617.

——AND SOSKICE，D. “The Economics of Sanctions，” Mimeo，Department of Economics，University of California，Berkeley，December，1976.

ARNOTT，R. AND STIGLITZ，J. “Equilibrium in Competitive Insurance Markets.” Mimeo，Department of Economics，Princeton University，1981.

ARROW，K. “Lecture Notes on Probability，” Harvard University，1970.

FARRELL，J. “A Sufficient Statistic Theorem，” Mimeo，Department of Economics，M. I. T.，1980.

FITZROY，F. “Contests，” Mimeo，International Institute of Management，West Berlin，1981.

GILBERT，R. AND STIGLITZ，J. “Entry，Equilibrium，and Welfare，” Mimeo，Department of Economics，Princeton University，1979.

GREEN，J. AND STOKEY，N. “A Comparison of Tournaments and Contests，” Discussion Paper 862. Harvard Institute of Economic Research，Harvard University，1981.

GROSSMAN，S. AND HART，O. “An Analysis of the Principal-Agent Problem，” Cambridge University Working Paper，1980.

HART，O. “The Market Mechanism as an Incentive Scheme，” Economic Theory Discussion Paper 53，University of Cambridge，1981.

HOLMSTROM，B. “On Incentives and Control in Organizations，” Ph. D. Dissertation，Stanford University，1977.

——，“Moral Hazard in Teams，” *Bell Journal of Economics*. Vol. 13，No. 2（Autumn 1982），pp，324 –340.

LAZEAR，E. AND ROSEN，S. “Rank Order Tournaments as Optimum Labor Contracts，” *Journal of Political Economy*，Vol. 89（1981），pp. 841 –864.

MIRRLEES, J. " Notes on Welfare Economics, Information, and Uncertainty " in M. S. Balch, D. L. McFadden, and S. Y. Wu, eds., *Essays on Economic Behavior under Uncertainty*, Amsterdam: North-Holland/American Elsevier, 1974.

——, "The Theory of Moral Hazard and Unobservable Behavior." Mimeo, Nuffield College, 1975.

MOOKHERJEE, D. " Optimal Incentive Contracts in Multiagent Situations," Ph. D. Dissertation, London School of Economics and Political Science, 1981.

NALEBUFF, B. "Prizes and Incentives," D. Phil. Thesis, Nuffield College, 1982.

——AND STIGLITZ, J. E. "Information, Competition and Markets," *American Economic Review* (May 1983).

——AND VARIAN, H. "Some Aspects of Risk Sharing in Nonclassical Environments," University of Michigan Working Paper, 1981.

O'KEEFE M., VISCUSI, K., AND ZECKHAUSER, R. " Economic Contests," Mimeo, Kennedy School of Government, Harvard University, 1982.

ROTHSCHILD, M. AND STIGLITZ, J. "Increasing Risk, I: A Definition," *Journal of Economic Theory*, Vol. 2 (1970), pp. 225 – 243.

STIGLITZ, J. "Incentives and Risk Sharing in Sharecropping," *Review of Economic Studies*, Vol. 41 (1974), pp. 219 – 256.

——, "Incentives, Risk, and Information: Notes towards a Theory of Hierarchy," *Bell Journal of Economics*, Vol. 6, No. 2 (Autumn 1975), pp. 552 – 579.

——, "Competition and Cooperation: Towards a General Theory of Compensation," Preliminary Work, Princeton University, 1980.

——, "Theory of Competition, Incentives, and Risks," Paper presented at the IEA Conference in New Developments in the Theory of Market Structures, Ottawa, May 1982, and forthcoming in the Proceedings, edited by J. E. Stiglitz and E. Mathewson.

——AND WEISS, A. Part I, Part II: "Credit Rationing in Markets with Imperfect Information, Part I." *American Economic Review* (June 1981); Part II published as Princeton University Econometric Research Program Research Memorandum 268, August 1980.

VARIAN, H, "Redistributive Taxation as a Social Insurance," *Journal of Public Economics*, Vol. 14 (1980), pp. 49 – 68.



信息与公司治理

信贷市场与资本控制*

过去在讨论资本市场的功能时，一般都认为资本市场能执行以下一些独特职能：在竞争性的资金使用者和用途之间配置稀缺的资本，[1] 为管理者提供信号以引导其做出正确的投资决策。[2] 在本次演讲中，我主要阐述资本市场另外一个相当不同的职能，我将其称为*资本控制*。[3]

传统的理论用一种拟人的方式来描述公司：它像一个理性的个体去做所有应该要做的事情，争取最大化公司股票的市场价值。在过去由小企业承担主要生产职能的时代，这种描述或许较好地反映了现实。每个小企业都在所有者的紧密控制下运营，而这些所有者又是在一心一意地追求财富，使自己强烈的欲望能得到满足。但是，现今多数生产都是由大公司完成的。在这些大公司中，每个股东都只能持小部分股份。因此，六十五年前弗兰克·奈特

---

* Credit Markets and the Control of Capital，本文是关于“货币、信贷与银行业演讲”。衷心感谢国家科学基金会所提供的资助。写作本文时，作者为斯坦福大学胡佛研究所的访问学者。作者与富兰克林·阿伦（Franklin Allen）、布鲁斯·格林瓦尔德（Bruce Greenwald）、山迪·格罗斯曼（Sandy Grossman）、奥利弗·哈特（Oliver Hart）、R. 吉尔伯特（R. Gilbert）、M. 斯皮格尔（M. Spiegel）和 A. 魏斯（A. Weiss）等人就此次演讲的主题进行了广泛的讨论，从中受益良多，在此谨致谢忱。约瑟夫·E. 斯蒂格利茨写作时是普林斯顿大学的经济学教授。“货币、信贷与银行业演讲”发表于1984年6月26日在拉斯维加斯召开的西部经济学会（the Western Economic Association）会议。

1 例如，在 Stiglitz（1982b）中，我就广泛地讨论了这种功能。同时参见 Stiglitz（1981）。

2 托宾的 q 值理论（Tobin's q theory）强调了股票市场的这一职能。

3 这并非意味着我们穷尽了信贷市场所执行的全部职能。例如，在另外一些文章中，我和格罗斯曼就强调，在将信息从不知情的投资者传递至知情的投资者那里，以及在汇总信息的过程中，价格发挥着重要的作用（Grossman and Stiglitz 1976，1980b）。

(Frank Knight）在其著作中十分关心的问题，即大公司的所有权与控制权的分离，业已成为现实。那些管理大公司的人能支配数额巨大的资本。他们若能很好地执行职能，经济就会繁荣，公司的股东也会因此而获利。假如他们不能充分地执行职能，或者假如他们将公司的资源用于满足自身目标，那么经济增长和股东利益都会受到损害。今天我想要阐述的问题是，在我们的社会中，到底使用哪些机制才能让我们确信，那些受托管理公司资源的人会很好地完成工作？究竟用哪些制度才能“控制”住那些支配公司资本的人？这些制度有效吗？能否使这些制度变得更有效？

在开始详细讨论之前，很有必要概括一下我将要提出的一些论点。在本文第一部分，我将会对传统的企业理论进行评述。这一理论的研究对象是由所有者管理的企业。传统的企业理论认为，当所有者自己不能管理企业时，管理者也会根据股东的利益来运营企业。银行就像铅笔或手纸的供应商一样。在企业的生产过程中，每个供应商都是一个必要的组成部分，因此银行也就没有任何理由要获得比其他投入品供应商更多的控制权。民粹主义者(Populist）认为，银行管理着整个国家。在19世纪晚期，这一观点是很普遍的。而在经济学（尤其是在制度经济学家）中，有一个延续甚久的传统。这一传统下的经济学认为，当银行不能管理企业时，管理者在运营企业时会根据自己的利益行事，而并不一定会依股东的利益去管理企业。现代经济学家在很大程度上摒弃了这一观点，他们认为有许多机制能确保，那些不依股东利益行事的管理者会被解雇。我情愿将其视为理论胜过现实的表现。而对于像我这样的理论家而言，这种胜利是一个很好的征兆。然而，它却更像是思维方式对理论和现实的一次胜利。事实上，在考虑了信息成本、风险厌恶以及管理活动的“公共物品”（public good）之特性后，经济理论就与民粹主义者的观点（我将用它来粗略地代表某些思想）更一致了，而与那些我将称其为19世纪的企业模型的观点就不太一致了。[4,5] 股东不能控制公司，

4 毋庸置疑，试图用一句话来概括广泛的理论，将会冒犯这些理论的支持者；而将这些理论与特定的时期联系起来，也可能会触犯经济思想史学家。让我感到甚为遗憾的是，限于篇幅，我无法将这些反方观点的某些歪曲之处也写下来。无论如何，这些有意的歪曲，确实能使我将要强调的独特的观点，得以更清晰地展现出来。

5 因此，我们得到的结论并不是特别新颖。相反，本文的目的只是为目前似乎不流行的理论奠定更坚实的基础。近来，又有许多研究重新关注此处所讨论的问题。这其中最引人注目的就是，最近在胡佛研究所召开的“公司与私有产权”研讨会（会议论文已发表在（*Journal of Law and Economics*）上）。当然，在那些论文中，也有许多文章的观点与在此所提到的观点并不相符。

而管理者也并一定就会依股东的利益行事。

尽管本文不会去证明任何命题，但我还是要用到信息经济学中最新的发展，尤其是在控制问题（通常将其称为“委托—代理”问题）上的发展。在控制问题中，一方试图通过间接的控制机制，引导另一方根据他的利益行事。[6]我们认为，若把下面将要讨论的问题称为多重委托—代理问题，也许更恰当一些。而对于多重委托—代理问题而言，纳什均衡在约束条件下是帕累托无效的（constrained Pareto inefficient）。[7]因此，我们在第二部分给出的公司模型，就是对此类一般性的多重委托—代理问题的一个解释。在第三部分，我们考察了资源配置中一些特别的偏差，而这些偏差就是由多重委托—代理问题造成的。第四部分提出了一些建议，以改革我们的制度。这些建议或许能减轻我们所指出的一些问题。

一、传统的观点

在原初的资本主义（primitive capitalism）模型中，每个企业都由唯一的所有者所控制。一般认为，我们所关心的问题并不会出现。这是因为，所有者—管理者获得其劳动成果，并承担他所犯的错误带来的成本，于是他就有最强的动力付出恰当的努力，去做出正确的决策。然而，只有在企业违约的概率为零时，这一观点才能成立。而不管是对19世纪的企业，还是对今天的公司而言，违约的概率为零这一假设都是不现实的。当企业违约时，导致企业违约的那些错误的部分成本，就不是由做出错误决策的主体来负责，而是要其债权人来承担了。这一认识严重地阻碍了有限责任公司的发展。[8]所

6 有关的参考文献和对这些理论的更为详尽的描述，参见本文第二部分。

7 在评价一个经济体的效率时，若用到了“约束条件下的帕累托最优”（constrained Pareto optimal）这个术语，那就表明，我们明确地考虑了信息成本。福利经济学基本定理最为重要的作用就是，它确定了使竞争性经济实现帕累托最优的一组特别假设。假如市场是不完善的（事实上确实如此），或假如信息是不完全的（事实上也确实如此），那么该经济一般不能实现约束条件下的帕累托最优（Greenwald and Stiglitz，1984）。尽管格林瓦尔德和斯蒂格利茨的研究表明，原则上讲，确实有一些对市场的干预措施（譬如，税收和补贴）能使每一个人的福利都得到改善。但是，一个合理的问题是，那些经由政治决策过程而实施的干预活动，是否确实都是帕累托改进的活动？对此，我们不予置评。然而，在最后一部分中，我们提出了一些改革的建议。这些建议也许能改进现状。但是，我们认为，并不一定就要用立法的形式来执行。尽管在某种程度上，立法会妨碍那些建议的引入，但是若能改变立法的话，那就最好不过了。

值得注意的是，我们在这里所关注的信息问题，还会以其他一些方式影响到竞争的结构。例如，这些信息问题自然会导致长期合作关系的形成（Stiglitz and Weiss，1983）。

8 对这一问题的讨论，参见 Allen（1980）。

有者并不会承担其行动的全部后果这一事实，有两个重要的含义：与承担行动的所有后果相比，所有者—管理者倾向于采取更大风险的行动；[9]贷款人也就倾向于采取各种方式去控制借款人的行动。后面我们会回到贷款人实施控制的机制上来。

正如我们已经指出的那样，用以刻画19世纪那些由所有者管理的企业的模型，与现代资本主义经济是非常不相称的。在现代资本主义经济中，部分管理者对公司也有责任感：他们之所以会最大化公司股票的市场价值，是因为他们认为一个称职的管理者应该这样做。他们所接受的教育都在讲，应该要去最大化公司股票的市场价值。管理者的选择程序也部分地体现了对接受过这方面教育的管理者的寻找。[10]但是，标准的经济学理论有一个很重要的组成部分，即每个人都是以自利的方式去行事。自利的非所有者—管理者为什么又会去最大化公司股票的市场价值呢?

对于这一问题，已经有三种解释：股东大会、接管以及“用美元投票”。我将会说明，为什么在这三种控制机制中，没有一种能非常有效地发挥作用。

1. 股东大会失灵

股东大会是一个会议场所。从严格的法律意义上讲，股东通过股东大会

---

9 若所有者—管理者是风险中性的，则（在有限责任下）他的支付函数为 $\max[R-(1+r)B, 0]$。其中，R 为项目收益，$(1+r)B$ 为公司应该支付给贷款人的数额（B 是债务额，$(1+r)$ 是每单位负债所带来的需要支付给贷款人的数额）。这是一个凸函数。因此所有者—管理者就会以一种风险爱好的方式行事。即使所有者—管理者是风险厌恶的，与不存在有限责任时他的行为相比，他仍然会以一种风险厌恶程度更低的方式（这是由 Diamond and Stiglitz，（1974）界定的）去行事（Stiglitz and Weiss，1981）。

10 说人类许多行为都是各种训练（conditioning）和（或）练习（training）的结果，这是没有多大争议的。绝大多数人之所以会去投票，只是因为他们认为这是公民的责任，而不是在对投票的预期成本和收益进行仔细计算后做出的决定。在绝大多数大学里，报酬结构几乎完全依赖于一些模糊不清的责任感。这种责任感要求每个人以最有效的方式，去执行各项应该要完成的任务。这么说并不意味着，就不会有（理性）自利的考虑了，只是表明我们不能简单地套用理性自利的原则来解释人类行为。

乔治·阿克洛夫（George Akerlof，1970）为对那些貌似非理性的行为，提供了一个富有洞察力的“理性的”解释。由于具有诚实等优点的人会在社会上得到较多的回报，因此父母就希望自己的孩子看起来具有诚实等优点。更理想的情况是，他们也许更希望自己的孩子表面上看起来很诚实，而实际上却不诚实。原因在于，这样他们就既可以得到那些能支配资源的职位，又可以将这些资源用于满足自己的目标。不幸的是，或者说幸运的是，当孩子们实际上并不诚实时，就很难让他们变得表面上看起来很诚实。假如这一判断成立的话，那么“次优的”策略就是将其培养成诚实的孩子。

就可以控制管理层。任何一位不采取行动去最大化公司股票的市场价值的管理者，股东都能够（而且原则上也应该）将其解雇。然而，在实践中，大量的证据表明，股东大会通常都是由管理层控制的，因此理论研究就需要考虑这个情形，即每个股东都认为自己无法对最终结果产生什么影响的情形。要获得信息以确定一个管理者是否为称职的管理者，以及要评价可供选择的管理团队，换句话说，要把票投好，总是要花费一定的成本。而且，这样做的收益又是那么微不足道。因此，就不会有股东为了把票投好而去花费资源。很久以前我们就认识到，运行良好的政府是一种公共物品。也正因为它是公共物品，所以人们就极少将资源花费在其上，以保证政府能很好地运行。但是，一家由公众持股的公司就像一个（地方）政府。[11]对于所有持有公司股票的人（或公司的债权人）而言，确保公司能良好地运行，就是一种公共物品。[12]

因此，如果说股东大会通常无法对管理层实施有效的控制，这一点儿都不奇怪；要是说股东大会能产生积极的效果，那倒令人奇怪了。[13]

2. 作为一种控制机制的接管失灵

接管提供了第二种控制机制。据说利用接管这一控制机制，资本市场可以确保不称职的管理者会被解雇。假如某些管理者不能让公司股票的市场价值最大化，那么就会有人来购买这家公司，并将公司策略改为价值最大化策略，而公司股票的市场价值的变动就是他施行良好管理的收益。然而，至少

---

11 斯蒂格利茨（Stiglitz，1977）进一步讨论了这一类比。

12 尽管一直以来，我们的学校在教导孩子们时都说，在大选中投票是一种美德，但是仍然很难看到以相同的方式教育孩子们：执行在股票上的投票权也是一种美德。

13 风险厌恶的投资者会充分地分散投资，以至于他在任何一个公司中的投资，相对于他的财富而言都是微不足道的。这一事实就进一步表明了，为什么理性的投资者都不应该为在股东大会上投好票而花费资源。

相反，如果投资者将很大部分财富投资到了一个公司的股票上，同时他们又获得了该公司很大部分股票的话，他们就能对结果产生一定的影响了。于是，对于他们而言，理性的决策就是，花费一定的资源确保公司能得到良好的运营。但是，这些投资者通常又会加入管理层，因此他们也就不会对管理层实施任何控制。至于为什么公司最初的所有者并不会完全分散投资，而要保留其所创立的公司很大部分的股票这一问题（当与不同公司的特征有关的信息不完全时，就会出现此问题），斯蒂格利茨（Stiglitz，1982b）给出了一个解释。

应该注意到，在某种程度上，这些最初的所有者确实是有动力去实施控制的。他们有动力去实施控制，以增加其自身的利益。他们的利益与小股东的利益并不一致（即使他们不以其他股东的利益为代价，来将资源用于满足自己的目标，譬如也许只是在风险态度上存在差异，这一结论仍然成立）。参见 Grossman and Stiglitz，1977，1980a）。

有四个原因会使得接管不能（或不太可能）成为一个有效的控制机制。[14]

首先，当一家公司运行得不太好时，要么是由于管理层不称职，要么是由于公司的资产实际上并不像它表面上那么好。[15]在这方面，与外部人相比，内部人（公司的管理者）可能知道得更清楚。因此，当那些在公司中拥有可控权益（a controlling interest）的人愿意出售其股票时，[16]这就表明试图接管该公司的个人或公司出价过高了。如果他们拒绝出售其股票，那就表明试图接管该公司的个人或公司出价过低了。因此，只有当投标的公司出价过高时，接管才能成功。[17]

当不对称信息使得问题产生后，无能的管理者就可以让有关信息不扩散出去。因此，就要花费一定的成本才能确定，经济体中哪些公司没有运营良好，并因此而低于其潜在价值。假定某一家公司分配了一些资源去审核许多公司，以便发现经营水平最差的那家公司。然后，它进行接管投标。这立刻就给了其他公司一个信号，它们会评价这个出价是否低估了目标公司的真实价值。如果那个出价确实低估了目标公司的真实价值，那么它们将会“投标”，击败发现目标公司的那家公司。在均衡状态下，这些“第二个”投标人的出现会使得他们的期望为零。但是，这也就意味着，发现目标公司的那

14 这些原因并非与所有情形都相关。

15 也就是说，给定公司过去的错误，现在公司也许已经运行得很好了，与其他任何人来做都一样好。

16 这实际上隐含地假设了，那些能决定接管投标结果（或对接管投标结果产生重要影响）的人，会根据全体股东的利益行事。当然，下面我们将会看到，通常他们都不会这么做。

17 与二级劳动力市场（secondary labor markets）就像是在雇佣“二手劳动力”（“used labor”）相似，接管就像是在购买“二手公司”（“used firms”）。在此应用阿克洛夫（Akerlof，1970）关于二手车市场萎缩的原创性思想，就会发现接管很像二手车市场的运行（Stiglitz，1975a；Greenwald，1984）。然而，与劳动市场或二手车市场相比，阿克洛夫的阐述甚至更适合这一市场。在劳动市场和二手车市场上，当不同的人具有不同的能力，而这些能力对于不同的企业又有特殊用途时，或者不同的人对二手车的质量有着不同的偏好时，交易可能就会发生。在接管中，只有运营公司资产的能力差异才是相关的（在理性预期条件下，仅有信念上的差异并不能带来交易（Stiglitz，1982b）。将这个一般性的结论应用到接管问题上的研究，Stiglitz and Hart（1981））。为了避免就那些对接管的效果进行评价的经验研究展开争论，我只是简单地指出，我尚未发现有足够说服力的证据能表明，接管能显著地提高被接管的公司的生产率。看起来更常见的情况是，若接管成功，则接管公司的股价会下降。如果真的如此，这就意味着，市场将接管的“成功”视为接管公司支付了过高价格的一个证据。同时，还提供了更深层次的证据，即公司（此处为接管公司）并不是由股东控制的。

即使有些公司对某一家公司的价值都有合理的意见，这些公司的数量相对而言也是很少的。相反，更常见的情况是，有些公司对目标公司的评价更高一些。因此，合适的模型不是标准的完全竞争模型，而是一个投标人数有限并且可能包含不对称信息的拍卖模型。

家公司的期望利润为负，因为它还花费资源去评价了许多运营得不是很差的公司。[18]

使接管机制变得无效的第三个原因与“公共物品”的观点有关。前面我们在讨论股东大会的无效性时，已经阐述了这种观点。假如接管成功，并且假设接管使得每股市场价上升了，则没有出售股票的那些股东就“搭便车”了。由于每一个小股东都认为（不管接管是否成功），它不会对结果产生任何影响，因此继续持有手中的股票符合其自身的利益。[19]当且仅当他认为，接管会成功，而且接管会降低其股票的市场价值时，他才会有动力卖出其股票。[20]因此，尽管价值降低型接管很容易会成功（存在一个理性预期均衡，在此均衡中，所有这一类的均衡都会成功），但是价值增加型的接管却很难成功。[21]

使接管机制变得无效的第四个原因是，当前在位的管理者通常都会采取策略型行动，以阻止接管。[22]尽管直到最近公众才开始关注，那些最引人注目的措施——金色降落伞；有条件的出售公司资产；如果接管成功，那么该

18　参见 Stiglitz（1975a）。一般而言，在做出接管投标决策，或决定是否应该评价一个接管投标以决定是否参与竞标时，各个公司都会采取混合策略。

19　如果有一些成熟的期权市场，那么从理论上推测应该就有，股东在分享所有收益的同时还能出售其股票。这些期权市场会导致一种证券的形成，这种证券等同于无投票权的股票。

20　应该要注意到，当仅有一个股东时，就不会出现这个问题。因此（忽略掉前面所讨论的不对称信息问题后），就可以认为目标公司处于连续拍卖之中。目标公司将会卖给那些认为自己能从公司资产中获得最高收益的人。公司资产对他（或她）的价值，与对出价第二高的竞标人的价值之差，就是出价最高的竞标人的租金。若竞标的人很多，而且当他们在管理的比较优势方面没有显著的差异时，租金基本上就等于零。尽管这种说法可以解释，在“原始的资本主义经济”中，公司资产得到了有效利用这一现象。但是，它却不能很好地解释现代资本主义经济中的情况：任何企业家都没有足够的资本，去获得任意一家大工业公司的所有股份。即便他可以成功地做到这一点，也很可能无法使其投资充分地分散，从而可以用风险中性的方式去行事。而且，如果他能成功地做到这一点，他也需要从金融机构借入大量的贷款，于是，金融机构就会对其行为施加一些重要的约束——这是本次演讲的主题。

格罗斯曼和哈特（Grossman and Hart, 1980）在解释接管机制的无效性时，就强调了这一点。因此，他们认为，使接管公司能得到补偿的唯一手段就是，通过转移公司资源来用于满足自己的目标。但是，如果那些能提高公司效率的人都这么做的话，他们就无法提高公司效率了。对于外部人而言，就很难判断目标公司是否得到了有效的管理。

21　很明显，这一分析依赖于试图进行接管的公司所面对的规则。如果接管公司可以发出收购要约，去购买51%的股份，接着还能采取行动去稀释小股东的利益，那么接管人就会很轻松了。但是，正如我们下面将会看到的那样，使价值增加型接管变得更为容易的规则，也可能会使价值降低型接管变得更加容易，还可能会产生均衡确实不存在这种情况。Stiglitz（1972）。

22　对这些策略性行动的讨论，参见 Cary（1969－70）。

项收购将会违反反托拉斯法。其他一些行动，比如含有违约后需要支付高额罚金这类条款的长期合约，会限制任何一家接管公司的行动，因此使接管变得更没有吸引力。[23]

3. 用美元投票

通常都假定还存在第三种控制机制：与梯伯特（Tiebout，1956）模型相类似，在地方公共物品理论中，每个人都可以“用脚投票”，而在股票市场上，每个人都可以用美元投票。那些不能有效地使用资源的公司，将无法筹集到额外的资金。然而，（与前面两种机制一样）这一机制也受到了重要的限制：从某种程度上讲，当公司只能从股票市场筹集资金时，它才能发挥作用，管理者对公司的现金流有相当大的自主权。对有些公司而言，它们有（超过其现金流的）很好的投资机会，因此当前最合意的行为就是去“贿赂”股东，以便未来可以筹集到额外的资金（而且，即使那样，投资机会“很好”也并非就意味着“有效的利用资源”）。但是，对于其他企业而言，它们的投资机会很差，因此在未来无法筹集到资金这一威胁，就不再是一个有效的控制机制了。

此外，我们注意到，当公司要回到市场来筹集额外的资金时，它们基本上都会求助于银行。这么做的原因有二。首先，发行股票通常会发送一个与“公司质量”有关的信号，并会导致股价的下跌（Greenwald，Stiglitz and Weiss，1984）。其次，与利用股票市场来筹资相比，通过银行筹集资金能更

23 接管不仅只有这些问题。如果接管人并没有购买目标公司的所有股票，那么，接管公司也许就会转移公司资源用于满足自己的目标。因此，在接管模型中，出价最高的公司并不一定就是那家能最有效地使用公司资源的公司，而可能会是那家能最有效地转移公司资源以满足自身目标的公司。虽然反对欺诈的法律并不能阻止这种偏离，但是确实会使某些公司（上游或下游的供应商或采购商）比其他公司更有优势。

但是，值得考虑的问题，远不止欺诈或将公司资源用于满足管理者个人的目标。假设在公司资源如何才能得到最好的使用这一问题上，总人口中有三类或三类以上的人的观点的可信度存在差异。那么，A 类型的人可以从 B 类型的人手中购买大量股票，并改变公司的生产计划。那些卖出股票的人——其中包括目前的所有者中对公司前景并不看好的那部分人——也许能稍微改善一下其境况；而那些没有出售股票的人，其境况可能会变差。不管当前的市场价格是上升还是下降，这都是正确的。B 类型的人认为，在最初的生产计划下，他们的期望收益会高于从其他资产中获得的收益（不过他们当中还有数量很少的一部分人，这些人是边际上的股票购买者。这部分人是另外一种类型的人，我们将其称为 C 类型的人）。同时，一旦 A 类型的人控制了公司的股票，则 B 类型的人就会出钱让大部分 A 类型的人放弃控制权。于是，就不存在均衡。这里分析隐含着不完全竞争这一假设：有些人能从其对特定公司的所有权中获得一些“剩余”；对该公司股票的需求曲线是向下倾斜的。有许多证据支持这种认为股票市场是“垄断竞争”市场的观点。参见 Stiglitz（1972a，1974，1975b）。

有效地控制资本。[24]这就是本文的主题。

4. 控制机制失灵的后果

我已经论述了，不管标准的“控制机制”在原初的资本主义经济中是否有效，它们在现代资本主义经济中，至多也就能发挥有限的作用。很难得到可以支持我的观点的经验证据：这种观点中至关重要的一点就是，外部人很难判断一家公司是否得到了有效的管理。尽管如此，依然有一些直接和间接的证据能支持我的观点。第一，有一种“广为人知的”技术，即考虑税收及避税行为。大量的证据表明，相当一部分公司并不会去使税收负担最小化（包括直接税，例如，公司税和间接税，例如，股东和债权人应该要支付的税收，对这些税收“悖论”的评论，参见 Stiglitz（1982a））。[25]第二，封闭型基金在很长一段时间内都以折价销售（这就意味着，基金经理随时都可以采用能提高基金市场价值的策略），以及接管后遇到的困难，都提供了坚实的证据（与此相类似的一种情况是，许多管理者都会抵制“资产剥离式接管”，即便资产剥离后的价值明显高于当前公司的价值）。第三，在最近一些接管大战中，管理者的行为与我们在此提出的观点更一致，而不太符合标准的观点。[26]

二、银行与资本控制

现在，简单明了地提出我们的问题：在资本主义经济中，我们如何才能确保那些受托管理公司资源的人，即大型公司的管理者，会有效地使用资

24 这并不是说，在资本主义经济中，就只有这一种控制机制。比如，有许多演化的观点表明，那些管理良好的公司将会存活下去，相反那些管理不善的公司就会被淘汰掉。对这些观点的批评，可以参见 Stiglitz（1974，1982a）。

25 有必要了解一下，在管理者控制的公司中出现这类看似反常的行为的比率，是否就大于在所有者管理的公司中出现此类貌似反常的行为的比率。

26 当然，我也意识到，对于真正信仰有效市场和有效公司的人而言，这些论述尚无完全的说服力。这就似乎跟托勒密的论述一样了，即每一个证据都与被我大致称为“传统的”理论相一致。本文的目标是，提供一个简单的可供选择的理论。这一理论与事实是一致的，并且在此理论中，事情实际上就是看起来那样的：那种看似与价值最大化不一致的行为，确实违反了价值最大化原则。

然而，也应该要注意到，价值降低型接管也可能会成功这一论断，可能就为管理者抵制接管的行为提供了一个合理的解释。他们这样做是符合股东利益的。

源。[27]我已经论述了，传统的控制机制是无效的。本文的中心思想就是，尽管法律将控制管理者的责任赋予股东，不过，从某种程度上讲，是银行和其他贷款人在控制管理者，而不是股东在控制管理者。但是，在阐述这一控制机制并讨论其后果之前，还需要详细地解释一下，我所理解的控制的含义。

就通常的用法而言，我们当中的绝大多数人都熟知控制的概念，但它的含义有时很模糊：我们会说对孩子失去了控制，不过接着又承认从来都没有控制过孩子；我们会问是谁在管理（或控制）大学，但是似乎又很难讲清楚权力到底来源于哪里。在我们所讨论的问题中，公司的管理者有责任管理好公司的资产，他们能以各种方式来利用公司的资产。我们说他们控制了这些资产的处置权。但是，他们的控制权也不是不受约束的。他们所管理的公司从银行和股东那里获得资源（资金）。在提供资金时，每个人都会对管理者施加一些约束条件。这些条件会限制管理者的行动，还会限制他的自由。[28]每一个人所施加的约束非常不一样。股东不能要求公司对其提供的资金支付一定的回报（他可以劝说另外一个人买下他的股票，不过，他一旦将资金交给了管理者，就不能迫使管理者将钱还给他）。他能在股东大会上执行投票权。贷款人会给贷款设定一个期限，期限届满时，他就可以回收贷款。如果管理者不按他的要求办事，贷款人就有权进行干预。干预的权利是由贷款合同和法令界定的。在其他一些经过严格界定的条件下，贷款人也有权进行干预。例如，贷款合同也许会对公司承担的项目的额外贷款进行限制等。[29]应该注意到，贷款人回收其款项的权力也是受到约束的。不过，毋庸置疑，与股东回收资金的权利相比，贷款人所受到的限制要少得多。

在贷款合约中，既有表达得很清楚的条款，又包含一些只可意会不能言传的条款。隐性合同（implicit contracts）是不能通过法律来执行的，只能用

---

27 正如我在前面所提到的那样，至关重要的问题并不是简单的欺诈。大致而言，欺诈还是可以直接通过法律来解决的。

有一种观点认为，管理者对公司的绩效影响几乎没有什么差异。经常被引用的证据就是，替换管理层并不能对公司绩效产生显著影响。其实，这种说法有点随意。这是因为管理者的选择过程是相当有效的，所以称职的管理者替换掉的也是称职的管理者，或者是因为管理者没有什么差异（另一方面，甚至很可能更合理的是，尽管无能的管理者明显会损害公司的利益，但在一定范围内他们似乎又很称职，因此他们对公司的损害会相对较少）。

28 这并不是对他的行动施加的唯一的约束。法律制度施加了其他一些约束，与公司签订了合同的买主和卖家进一步施加了许多约束。

29 在有些情况下，法律会对贷款合同所设定的许多干预进行限制。如果贷款人对公司管理干预过多，那么他也许会失去法律赋予贷款人的那些优势。

其他的方式来执行：若借款人不遵守隐性合同中的条款，就可能无法将借款展期（同时，在所有意会的条款都得到遵守的情况下，如果贷款人不给贷款展期，那么他就会发现，可能很难找到愿意从他这里借款的人）。[30]债权人和股东对管理者的控制，不但可以通过直接的方式（比如，对管理者可能会采取的行动加以约束）来实现，而且还可以通过间接的方式（比如，设计合适的薪酬结构，以诱使管理者采取（更）符合债权人和股东利益的行动）来实现。

因此，贷款人与股东通过显性合同和隐性合同，对管理者施加的直接控制和间接控制都是不完全的。贷款人通过合同的正式条款以及拒绝给贷款展期，来对管理者实施控制；而股东则是通过投票和拒绝提供额外的资金来实施控制。管理者的激励既会受到显性的薪酬的影响，又会受到其他隐性的惩罚措施的影响。显性的薪酬是指，由其他可能聘用他们的公司所提供的报酬；而隐性的惩罚措施则是指，其他公司在对待那些被自己公司解聘（或公司破产）的管理者时所采取的手段。所有这些报酬、惩罚措施以及决定管理者行动的限制性措施，都不是由某一个人所能确定的，甚至都不能由一家公司的所有者和贷款人确定，而是由整个市场确定的。[31]没有任何一个人能单独控制管理者。这是因为，要有很多人和许多机构才能影响管理者的行为。虽然早期的一些文献[32]试图将管理者视为债权人（“委托人”）的代理人，但是，在债权人要设计一份有效的激励合约给代理人的条件下，更合适的模型是多重委托—代理模型。[33]在多重委托—代理模型中，每个委托人都

30 就像劳动合同一样，显性合同并不一定就比隐性合同好。显性合同的执行不但要求违约行为是可以观测的，而且还要求违约行为是可以由第三方（法院）证实。隐性合同的执行就不需要这些。另一方面，虽然一期的显性合同是可以执行的，但是，对于隐性合同而言，要想使其是可以执行的，就必须要建立起长期关系。

31 因此，如果借款人能很简单地转向其他贷款人获得贷款，那么某个债权人说要中止贷款的威胁，就起不了什么作用。如果工人能很简单地转向其他雇主并获得相同的工资，那么某一个雇主说要解雇工人的威胁，也起不了多大的作用。只有当这些威胁都有效时，或者说，只有当工人被解雇以及公司被通常会提供贷款的银行中止信贷都会带来损失时，才会存在市场均衡。对具有这种性质的均衡的分析，参见 Stiglitz and Weiss（1983）；Shapiro and Stiglitz（1984）。

32 这种观点最初是由 Stiglitz（1974）提出来的。其中，地主的问题与资本所有者的问题相类似。地主要诱导他所雇佣的工人努力工作，而资本所有者则要诱导他所雇佣的管理者努力工作（Ross，1973）。自那以后，大量的文献都在研究委托—代理问题。这些文献不仅强调要诱导合意的努力水平这个问题，而且强调要确保管理者只能承担适当的风险这个问题。

33 尽管应该要强调，在这里，“有效的”这个词只表明激励合约是局部有效的。也就是说，当价格等因素给定时，在市场上，若不使代理人的状况变差，委托人的状况就不能变得更好。然而，市场均衡通常都不是帕累托有效的。参见 Arnott and Stiglitz（1983a）；Greenwald and Stiglitz（1984）。

只能设定合同的部分条款。虽然这个问题在我们的经济中普遍存在，但是却很少会有人关注它。

我再举一个例子。其中，每个委托人自然会有一组控制权。以一个分成制经济为例，佃户要生产出产品，就同时需要资金和土地。假定其中一个人拥有资金，而另外一个人拥有土地。每一个人都与佃户签订一份合同，规定他提供其资源后能获得多少补偿。当然，佃户的行为会受到两份合同的影响。也就是说，在两份合同之间存在外部性：佃户若有巨额未清偿债务（当存在抵押劳动（bonded labor）条款时），与其没有这些债务时的情况相比，他所承担的风险更低，并且会更加努力地工作。[34]因此，他就有动力去将这些外部性内部化。例如，对于地主而言，他会同时提供贷款给佃户，或者至少会要求佃户披露其与他人签订的贷款合同。[35]但是，有许多原因使得完全的内部化通常是不可能的。[36]在这些情形下，任何一个委托人都无法控制代理人。我认为，在大的股份公司中出现的就是这种情形，即没有任何一个委托人能控制管理者。

在多人实施控制的情况下，我们就会问，谁能实施有效的控制？或者说，在实施控制时，谁最为重要？我想这是一个有意义的问题，但是却很难对其进行定量分析：显然，即使 A 能够施加的约束和他能获得的报酬，B 都能做得到，仍然会出现 A 比 B 更能有效地控制管理者这种情况，反之亦然。不过，通常都是一些 A 能做到而 B 做不到的事情，或者相反。一些股东能做而贷款人不能做的事情（他们可以在股东大会上投票），而贷款人也能做一些股东不能做的事情。这不是那种能精确定义的情形（根据我的目的，这也没有必要）。希望我将要给出的直观的概念，暂时能满足分析的需要。

在决定谁能实施有效的控制时，与之相关的不但有干预的权力及方式，

34 外部性是相互的：租佃合同条款也会影响贷款违约的概率。对这些问题更详细的讨论，参见 Braverman and Stiglitz（1982）。

35 对委托—代理（道德风险）问题背景下的外部性问题更加一般的讨论，参见 Arnott and Stiglitz（1983b）。对多重委托—代理问题有相当精彩的讨论，参见 Bernheim and Whinston（1984）。

36 有限的信息也许会使“排他性”合同不可能得到执行。另一方面，一个委托人也许无法提供代理人所需要的全部资源。因此，地主也许是没有充足的资本，或许是没有佃户所需要的耕牛。当然，他也可以（同他的佃户一样）从其他一些资源提供者那里获得这些资源，但是，这肯定又会产生新的委托—代理问题。

事实上，可能还存在进一步的限制。如果委托人为代理人提供其所需的全部资源，那么委托人的福利就严重依赖于代理人的行为。这是因为，风险并没有得到充分的分散。

而且还有激励：（从“干预者”的角度看）如果干预能使代理人的行为有所改正，那么干预就是合意的。为了确定这一点，干预者就需要了解，管理者目前的行动以及可供替代的选择。而了解这些信息是有成本的。

此时，我们就面临着基本的困境：如果管理者能从一个供应者那里获得所有资源，那么该供应者就有动力去收集信息，以确使管理者会按照供应者的利益行事。但是，供应者此时就要承担相当大的风险。这是因为，在此情形下，他不能充分地分散风险。另一方面，任何一个股东或小贷款人都不会去关注公司的绩效，其原因在于，他的行动带来了一些收益，同时给所有其他处于相同地位的供应者也带来了收益。这就是我们前面讨论过的“搭便车”问题。

因此，尽管贷款人和股东都有一定的权力去控制管理者，但是，从个人的角度看，他们都没有动力去实施这些权力。于是，管理者就未得到有效的控制。

如果这就是故事的结局，那么大规模生产的资本主义经济之前景，就会非常黯淡。假如管理者不能得到有效的控制，那么就不会有人将那些发展现代工业所必需的资金交给管理者。然而，还有三个控制机制，尽管它们的运行尚不完美，但是也足以使许多人愿意将资金交给其他人来管理了。

（1）银行。我猜想，在这三个控制机制中，最重要的就是银行（包括贷款银行和投资银行，投资银行能帮助公司筹资）。在公司中，银行通常处于很重要的地位。贷款合同的本质，会促使银行在不承担不当风险时去控制管理者。同时，贷款合同的本质，还会使银行将注意力集中在与下列问题有关的信息的收集上：跟违约概率相关的问题，以及低收益状态下公司的资本净值是多少的问题。银行既不需要考虑当公司最繁荣时它们的境况会有多好，也不需要考虑当管理者像强盗那样卷款潜逃的概率有多大。

由于银行关注的是小概率事件，同时由于管理者的激励结构鼓励管理者避免这些事件，因此银行实施有效控制的收益是有限的。[37]稍后我们再回到这一问题上来。

这种论述表明，对于贷款人而言，如果只有一个贷款人，那么就可以解决“管理的公共物品”问题。但事情并非总是如此：经常会有银团贷款，在银团贷款中，有一家银行会担当“领导”角色，并承担确保借款人得到

37　除了避免损失惨重的结果之外，担保既能降低银行的风险，又会削弱银行实施严密控制的动力。

了有效控制（从贷款人的角度看）的责任。[38]此时，就要通过声誉来解决管理的公共物品问题了：确保按照牵头银行的利益行事的贷款是优质贷款，这符合每一家银行的利益，以免其他银行拒绝参与由牵头银行组织的银团贷款。[39]

（2）**集中的股票所有权**。在债务这一情形中，可以通过由一个供应者提供所有资源（或者由几个供应者合作行动）来解决问题。但是，在股权分散的大公司中，股东的数目非常大，而且到底哪些人是股东通常也不清楚，因此上述机制就不能发挥作用。不过，如果只有少数几个股东，而且每个股东都与公司休戚相关，那么每个股东就会有足够的动力去控制管理者，于是他们也就会花费足够的资源去获取信息，利用这些信息，他们就能对管理者实施有效的控制（小股东仍然会搭大股东的便车；标准的分析表明，这会使得那种用于控制管理者的资源是不充足的，至少从股东这个整体的角度看是这样的）。这也是有成本的，即这些股东的风险分散程度受到了限制。此外，这些股东的利益，也许不会与小股东的利益十分一致。[40]从理论上推测，对于那些愿意承担有限风险分散的股东而言，他们必须要得到补偿，例如，通过转移部分公司资源来满足自己的目标（比如，为公司董事会提供很高

---

38 通常很难区分激励问题和甄别问题：牵头银行也有责任确定哪些借款人有资格获得贷款。

39 最近，许多研究都在讨论在多时期博弈中合作均衡的实施问题。尽管绝大多数研究都假设没有贴现（因此，对我们的分析而言，这些研究的用处非常有限），但是，近来也有几项研究开始考虑贴现了。尤其要参见 Abreu（1983）。

正如通常所分析的那样，在一时期模型中，每一家银行都可以从欺骗中获利：参与银团贷款后，也不需花费资源去确保牵头的贷款人恰当地使用资金。在多时期模型（multiperiod context）中，这就不正确了。一家“失去声誉的”且不能组织银团贷款的银行，就会承当更大的风险。可以将这种“风险贴水”视为对其不去施加控制的罚金。应该要注意到，与此相关的银行数目是充分少的，因此每家银行都可以评价其他银行的声誉。

在投资银行通过债券来为公司筹资时，也有一个类似的声誉机制在发挥作用。投资银行同意作为债券持有人的“受托人”去实施一定的控制。然而，当投资银行通过发行股票来为公司筹集资金时，这种观点就不适用了。股票发行完成之后，在确保公司的管理者恰当使用资源的过程中，投资银行通常都不能发挥任何作用。投资银行的责任，仅限于甄别那些通过股票来筹资的公司的质量这一方面（对风险资本和私募资本而言，这种观点就不太正确了）。

40 一个更富有的人拥有一家公司更多的股票这一事实，并不一定意味着，他就有更大的动力去获取信息和实施控制。原因在于，如果他的时间价值也同比例上升，那么获取信息的成本也会随着收益的增加而增加。

的费用)。[41]

(3) *管理者的声誉*。第三种控制机制是,管理者对其良好声誉的考虑。在管理者所服务的公司内外,都会有人对他的行为进行评价。如果他表现得"很好",那么就很有可能会得到提拔,或者是被其他的公司挖走。如果其他人认为,他的行事方式"很不好",那么最好的情况就是,他的外部机会就会少很多,而最差的情况则是,公司破产,债权人进行干预,他自己被解雇。他一直都面临着内部"政变"的威胁:下属或外部人会试图让董事会相信,他们可以把管理工作做得更好。[42]

当然,管理者不但会受到其声誉的影响,而且还会受到薪酬明细表的影响。例如,有人用这一观点来解释股票期权。最近,拉里逊(Larson,1984)对那些实施了一段时间的股票期权激励计划进行了研究。非常有趣的是,研究结果表明,那些实施了这一计划的公司的市场表现,并不比作为一个整体的市场的表现要好(以股票市场价值为标准)。在实施了管理者激

41 如果此类股东只能得到这点补偿,那么对管理这种公共物品的支付就很少了。不幸的是,在此处讨论的其他两种控制机制不能发挥作用的地方,控股股东和管理者在合作将公司资源转而用于满足自身目标这方面,似乎还有相当大的自由决断权。

应该要注意到,还有其他一些原因会导致有限的风险分散:公司的创始人会通过保留相当多的股份来向外界传递一个信号,即他对公司的发展很有信心(Stiglitz,1982b)。由创始人控制的公司,是本文第一部分讨论的原初的资本主义经济,与此处讨论的发达的资本主义经济的混合物。

就像公司的创始人会通过保留相当比例的股票,表示其对公司的发展的信心(并因此而提升公司股票的市场价值)一样,他也有动力去制定公司章程,以此为后来的管理者提供"良好的激励"。从理论上推测,股票市场会通过提升公司股票当前的市场价值,来回报他这一行为。尽管这种理论分析很明显是对的,但是通常会出现这种情况,即公司章程制定后,公司创始人仍然不想退休。也经常会出现另外一种情况,即创始人与公司签订一份长期合同,以保证他能在很长一段时期内,都可以为公司提供一定的服务。因此,设计一份能激励未来有潜力的管理者的合同这种行为,在10年、20年甚至30年内的贴现值,很可能非常少。最好是建议创始人将其注意力转移到其他问题上去。

42 事实上,董事会中通常会有几名来自管理层的人,因此董事会不会不知道,不同的管理者和董事的激励是不同的,他们并不是像一个团队那样行事。很明显地,公司的董事长要解聘首席执行官,那是要冒很大风险的。但这种"宫廷改变"并不经常发行。

励计划的公司中，几乎就没有能与经济理论的预测粗略地保持一致的公司。[43]

三、控制机制中的偏差

当股东不能控制公司时，我们提出了几个可供选择的控制机制。管理者会对这些控制采取应对措施。因此，我想要分析的是，公司不最大化其股票市场价值这一问题。[44]在下一部分，我将会讨论，是否有另外一些制度安排可以确保，资本能得到有效的使用。

在前面所描述的几种控制机制中，都存在一些系统性偏差。贷款人只关心收益分布尾部的最低部分。因此，他们会要求公司只做那些相对而言风险很小（最低的尾部）的项目，即使这样做的期望收益很低（或者，他们也许会通过一些间接控制机制、贷款合同中的控制条款、贷款展期的条件、他们所需要的贷款担保或者他们所收取的利息，来诱使公司去做那些风险很小的项目）。为了降低违约风险，他们需要确保管理者不会卷款潜逃。这就为所有投资者提供了一种公共物品。不过，从其他的角度看，银行的利益和股东的利益是相互冲突的。由于银行可以撤回资金，因此它就能在实施控制时

---

43 最近，汤普森—拉莫—伍尔德里奇公司（TRW）制订了一项管理者薪酬计划。这是一个引人注目的例外。股票期权计划，并不是将管理者的绩效与其他人在可比环境下绩效相比较后，再决定管理者的报酬，而是根据公司股票的市场表现来决定管理者报酬。为什么管理者一定要承担这种风险？在原则上，这么做有一个好处：管理者会更加关注公司收益与公司股票的市场表现的相关关系。很少有证据表明，在采用这些计划时，这是一个重要的需要考虑的因素。而且，也很少有证据表明，会有一些相当重要的结果出现。确实，如果管理者会更加关注公司收益与公司股票的市场表现的相关关系，那么公司与管理者所签订的合同，就应该对管理者购买指数期权或其他公司股票的行为进行限制（现今，因为期货市场的存在，所以管理者也许可以将自己所承担的市场风险进行转移。这些市场风险与他所获得的期权有关）。

与税收优势有关的标准的论述是不合理的。事实上，恰恰相反，即使对管理者的支付要取决于公司的绩效，最好的办法是直接支付奖金给管理者，而不是通过股票期权来对管理者进行支付（Stiglitz，1982a）（最近，当我在一个研讨会上提出这一观点时，与会的许多为美国的大公司设计管理者薪酬计划的人，没有人提出什么反对意见）。

注意到下面这一点，也是很有意思的：绝大多数的股票期权计划实施后，管理者似乎也没有承担更多的风险。原因在于，（据说可以为股东节省大量成本的）股票期权计划，只是作为管理者的工资结构的一个补充。对以激励结构为基础的相对绩效的讨论，参见 Nalebuff and Stiglitz（1983）。

44 这一问题本身并不意味着，经济就不是约束条件下的帕累托有效的。确实并不能像不存在信息成本那样，有效地使用资源，但是这个说法没有多大意义。众所周知，当存在委托—代理（激励）问题时，经济并不一定就是约束条件下的帕累托无效的（Arnott and Stiglitz，1983b；Greenwald and Stiglitz，1984）。

发挥作用。原因在于，管理者会对银行的要求做出反应。[45]

管理者对自身声誉的关心，以及管理者的认识（即对管理者的评价通常都是（隐性地）以他们的相对绩效为基础的）都意味着，他们会采取“从众行为”。如果按照惯例，公司应该支付股利，那么管理者就会安排公司支付股利，而不管这是否符合股东的利益。大型股份公司的任何一位管理者，若想对公司股东解释股利悖论，则大家都认为他不太正常，而且其未来的职业前景也会变差。[46]管理者也许会要求公司，按照惯例为他们支付报酬，而不考虑管理者的时间的机会成本，也不考虑基本上类似的管理者能获得的收益。尽管在使汽车产业的盈利能力恢复后，汽车产业中的管理者很快就会宣称，他们应该因此而获得更多的报酬（全然不顾这些奖金会对劳工关系产生严重的负面影响），但是，当他们使汽车产业遭受空前的损失时，他们就不会很快受到责备（并削减相应的报酬）。[47]管理者（最多）只是公司资产的临时受托人而已，而且他们的报酬结构在很大程度上都是由当前的收益决定的，因此，很少会有公司将十年或二十年的绩效作为基础，来决定管理者的报酬。即便有公司会这么做，也几乎不会要求管理者，在任职届满后很长一段时间内，还保留公司的股票。然而，许多长期投资的收益，要到未来的某个时间才能实现。在《就业、利息和货币通论》中，凯恩斯认为，股票市场中的投资者只关心短期收益，而不关心长期的回报。今天，越来越多的人用类似的观点来反对美国的大企业中的管理者：这些企业的领导都是财务专家，而不是生产专家。他们的工作就是分配资金，而且他们的观点与凯恩斯所讲的投资者并没有什么区别：他们希望，能在其他投资者发现某些资产被低估之前，就找到那些资产。这样，他们就可以获得短期资本收益。他们的行为并不令人惊讶。但问题是，有什么东西能激励他们去关心公司的长期

---

45 贷款的期限给贷款人提出了一个有趣的问题。一方面，短期贷款给了贷款人随时收回贷款（或进行干预）的权利，但是贷款人不能对借款人进行连续的监督，而且在贷款人进行干预期间，借款人也许会采取行动，以使得向贷款人偿付变得不可能。因此，他们就会威胁贷款人：要么对贷款合同展期，要么就承担公司破产所带来的风险。更有效的一种办法是，贷款人承诺在一段较长的时间内，为公司提供资金，并放弃某些干预的权利（Stiglitz and Weiss，1981）。

46 从理论上推测，当管理者进入公司的晋级序列后，他就能识别各种倾向，而且其进一步的发展也会碰到一些困难。

47 只要公司能不断使利润永创新高，贷款人就不会对这些支付进行限制；股东想对这些支付进行限制。但是，正如我们在前文所讨论的那样，股东无法对管理者实施有效的控制。

前景或经济体的生产率？[48]

由于小股东不能直接地对管理者实施控制，因此他们就必须依赖于其他一些控制机制。正如我们已经表明的那样，大股东也许有动力去实施控制，但是他们的利益与小股东的利益并不十分一致。类似地，尽管小股东也许会认识到，贷款人的利益与小股东的利益并不一致，但是他们知道，贷款人使管理者不能卷款潜逃，而这会在一定程度上确保其投资的安全性。同时，小股东还知道，管理者的收入通常与公司的经营成果相关，于是他们的激励至少有一部分是一致的。

四、制度改革

如果我们的论述是正确的，那么它们就表明：对传统的资本主义经济运行的认识，客气一点说，它过于简单化了；不客气地说，它是错误的。管理者拥有相当大的自由决断权。原因在于，控制机制都不完善。尽管更早的伯利—米恩斯（Berle-Means）认为管理者控制了公司的观点，以及马奇—西蒙（March-Simon）认为管理者只是寻求简单的满足的观点都不是很完善。但是，与那种认为公司会最大化其股票市场价值的极端观点相比，这两种观点也许更精确地描述了公司的行为。管理者面临由很多机构设置的诸多约束和激励。那些由分散的股东持股的大公司中的管理者，肯定会以不同于所有者—管理者的方式行事。我们已经指出了一些偏差：贷款人更关心如何避免公司违约，而不是使收益最大化；而管理者更关心短期的盈利能力，而不关注公司长期的盈利前景。

因此，问题就是，有没有这样一些制度改革？即便它们不能确保完全的

48 应该将他们的作用，与原初资本主义经济中的所有者—管理者的作用进行比较。所有者—管理者想建立起自己的王朝的雄心壮志就包括，将他的公司留给自己的继承人。这些人就不会每天都在想公司能卖什么呢？而是去关心公司的长期前景。

能表明管理者特别关心公司短期前景的证据，主要是所谓的税收悖论。比如，在通货膨胀时，广泛使用“先进先出法”（FIFO）这一存货会计核算法，以及在转而采用加速折旧法时，广泛采用很长的调整期。对这些现象标准的解释是，如果管理者确实做了这些改变，那么也是因为他们认识到，股票市场会对会计利润的下降做出错误的判断，并因此使公司股票的市场价格下降。即使公司实际的（税后）利润增加，股价仍然可能会下跌（如果投资者天真地，坚持是以当期报告的利润为基础，那么公司股票的市场价值在未来终究会上升的）。

显然，如果当期的市场价格精确地反映了公司的长期盈利前景，那么在最大化当期市场价值与最大化长期市场价值之间，就不存在任何差异。当风险和期货市场不完善，而且投资者无法完全了解公司的活动时，就没有理由相信，最大化当期市场价值仍然等同于最大化长期市场价值。前面所说的税收悖论就是这种情况。

效率，至少也能改进一下现状。

在给出一些初步的建议之前，重新叙述一遍我所注意到的问题。首先，绝大多数大型工业企业所需要的资源，超出了单个人的能力范围；即使某个人有这些资源，他所承担的风险，也超过了他所能承担的范围。概括地说，在资本主义经济中，大致有两种（有时也会有更多种）类型的资金供应者。一类（股东）有许多参与者。对于每个股东而言，公司的管理都是一种公共物品。由于实施有效的控制是有成本的，所以每个股东都想"搭便车"。结果就是，尽管名义上股东拥有控制权，实际上却不能有效地实施。另一类（贷款人）的参与人数较少，而且制度安排所提供的控制机制，也能减轻"搭便车"问题。每一类资金供应者都有其自身的利益，也掌握了有效的控制手段。这就好像是一个存在多个委托人/一个代理人（或多个代理人）的问题，而对于此类问题（即使忽略掉"搭便车"问题），纳什均衡也是无效率的。

因此，我们的目标就是，找到一些制度安排，以减少搭便车问题，让那些信息成本相对较低的人加强控制，并为那些能进行控制的人提供激励结构，以使他们在实施控制时将所有相关各方的利益都考虑进去。

在此，我想提出两个初步的建议。每个建议可能都有点问题。虽然我不能确信这些改革都是合意的，但是，我坚信在我们的经济中，制度创新的空间非常大：萎靡不振的经济和我们的分析都表明，我们不能用19世纪的思维方式，来看待今天的经济是如何组织的。

两个建议都是以监督成本是很重要的这一认识为基础的。也就是说，要想实施有效的控制，就必须要让那些能进行控制的人有很强的激励去做这件事。有两种机构具有这种激励：银行和（作为工人代表的）工会。[49]银行执行其控制职能的关键问题就在于，它们对收益分布的更低的尾部给予了过多的关注。让银行（或者，也许更好的办法是，让一个相关但独立的分支机构）拥有公司的股票，也许可以减轻这一问题。而且，应该要让银行在对管理层实施控制时，不但要关注收益分布的最低的尾部，而且要关注收益分

49　如果劳动市场是完全竞争的，而且个人在变换工作时没有相关成本，那么工人几乎就不会关心公司的前景。但是，这两个假设都是不现实的。

布的均值。[50]

另一个需要加强控制职能的主要机构是工会。实质上，它们有一个其他机构都没有的优势：它们（或它们的成员）对公司的日常运转十分清楚，因此对于它们而言，获取与公司有关的（特定类型的）信息的成本，很有可能要低于其他机构花费的成本。它们还有第二种优势：与银行一样，它们非常希望公司能存活下去。从整体上讲，在公司中，全部工人可能有最大的不可分散的利益。[51]已经有越来越多的人意识到，那种认为管理层与工人的关系是对抗性关系的传统观点是不恰当的：它并不是一个零和博弈。我们已经论述了，工人（工会）既有信息基础，又有动力去实施监督和控制。但是，他们也只会根据其自身利益行事。巨额退休基金的发展提供了一个工具，利用这一工具就可以使各方利益得到更多的兼顾：如果将退休基金中的相当一部分投资在自己的公司里，他们就会有动力确保管理者采取的行动，同时还符合股东的利益。在缩小工人利益与股东利益之间的差异这方面，员

50 这一提议并不是没有问题的：银行知道它的存款人都会得到联邦存款保险公司（FDIC）的保险，因此，会为了使其下属的机构获得更高的收益，而将存款人暴露在更大的风险之下。

最近，有些大银行承担了过量的相互关联的风险（石油和天然气贷款，对欠发达国家的贷款）。在它们身上发生的这些事情提出了一些问题。这些问题与银行在此处所讨论的条件下，执行其控制职能的能力有关。这表明，给定联邦存款保险公司提供的（显性或隐性的）保险，从所有者的期望收益这一角度看，或者从银行的角度看，这些决策也许确实就是最好的决策。或者说，无论如何，这些决策都是信贷员所面对的激励结构的自然结果。

我认为，只要产业要面对国际竞争压力，并且联邦贸易委员会和司法部能保持对反竞争行为的监督，那么认为银行会利用它们的地位去限制竞争的看法，就有点夸大其词了。此外，我们并不是说要将所有的控制权都交给这些银行业的机构，而是说要加强其控制职能，以使其能更多地考虑股东的利益。

在整个讨论中，我一直都假定，银行关心的是如何才能避免公司违约。尽管银行应该关心这个问题，但是在谈到这一问题时，经常会有人问我：银行是否就不会面临类似的委托—代理问题？我们怎样才能知道银行信贷员不会追求其他目标？而且，如果他们不会那么做的话，我们就不需要将关注的焦点简单地从控制一些管理者（公司的管理者）转向控制另一些管理者（银行的管理者）吗？为了回答这些问题，我给出两点评论。首先，我并没有说管理者是完全得不到控制的，而只是说他们有许多自由决断权。同时，我认为，银行的激励结构（支付）上的变化，至少会在一定程度上反映在银行的管理者所面对的激励结构中。其次，我并不是说，应该赋予更多的控制权给这些银行（银行的管理者），而是说我们应该考虑一些可供选择的制度安排，这些制度安排会使银行采取更符合股东利益的行动。

51 公司一旦破产，工人的损失（包括失业时的工资损失和特定培训收益的损失）也许会超过资金供应者的损失。此外，从个体的角度看，与资金供应者的损失占其财富的比例相比，工人的损失占其财富收入的比例可能更高一些。

工持股计划也许可以走得更远了。[52]

长期以来，有这样一个传统：工人应该进行劳动，管理者负责管理活动。在原初的资本主义经济中，这种职能的分离也许是合适的。但是，正如我们所论述的那样，现代股份公司的问题，已经不能简化为一个委托人/多个代理人这个问题了。相反，应该是多个委托人/多个代理人问题，而对于此类问题，纳什均衡通常都是无效率的。假定每个人都拥有一些控制权，那么每一个参与人（或每一类参与人）都追求自己的利益，会使最终的结果不是帕累托有效的。每一群人的行动都会对所有其他人产生重要影响（存在重要的外部性）。此外，我们已经多次强调，要注意到良好的管理具有公共物品的性质：对于任何一类资金供应者中的少数人而言，花费资源确保管理者会追求自己那一类资金供应者的利益，是不符合其自身利益的（而且在某种程度上，那些有利于某一类人的行动，会对其他人的利益产生溢出效应，于是每一类人提供的“监督”服务都很少）。

我们的问题是设计出一些制度结构，从而将这些外部性的一部分内部化，而且从那些能最有效地获得信息实施控制的人那里获益，并改善（显然是不能消除）部分“搭便车”的问题，“搭便车”问题根植于良好管理的维护之中。我们在此所提出的两个建议就是希望能达到这些目标。这两个建议能否扬名于世，或者是否会被其他更大的负面效应所抵消，还要看情况的发展。[53]

---

52 这一提议也不是没有问题的。从某种程度上讲，工人将其储蓄（退休金）投资于公司，那么与他们所能承担的风险相比，此时他们所承担的风险更大。

此外，工会是否会按照工人的利益去行事通常也不清楚。可以将工会看成是为了替工人解决公共物品问题而采取的一种手段。但是，与任何此类机构都一样，在当选官员的自身利益与他们应该为之服务的人的利益之间，可能会存在显著的差异。

53 我想，还有其他一些改革措施也许能改善现实情况。例如，可以通过将管理者的部分报酬的基础，定在公司在未来五年、十年或十五年的绩效上，以此来激励管理者考虑公司的长期利益。

鼓励股东长期持有公司股票的税法，尽管会削弱股票的流动性，但是却可以激励股东更加关注公司的长期盈利能力。

当股东利益与管理者利益的冲突看起来非常明显时（像有些金色降落伞），对某些行动施加一定的限制，可能会产生有利的效应。不过，从两者的利益差异中最明显的问题这方面看，它们很可能也只是使问题得到了一点缓解而已。根本性的问题依然存在。

还可以利用一些组织形式来从很多人手中筹集大量资金，并减轻公共物品问题。例如，假设限定每家公司最多只能有十个股东。然后，这些公司可以对更大的公司投资（依然限定每家公司最多只能有十个投资者，不过此处的投资者是“公司”，而不是个人）。在每一个阶段，投资者的数量都相当少，因此可以极大地削弱公共物品问题。这个制度唯一的缺点就是，存在一条委托—代理关系“链”。与这类链条有关的无效性，是当前研究的一个主题。

最后，希望我的分析能让大家认识到，用19世纪的所有者—管理者资本主义模型来描述20世纪的资本主义经济是不合适的。关于是谁在控制公司这一问题，并没有简单的答案。不过，有一点是很清楚的，就是小股东无法实施有效的控制，但银行却能对管理者实施更有效的控制，而且从某种程度上讲，当银行实施控制时，在它们的决策中也存在重要的偏差。我们的体制可能运行得很好，不过我想应该还有改进的空间：社会创新与技术创新一样重要。[54]

参考文献

Abreu, Dilip. "Repeated Games with Discounting: A General Theory and an Application to Oligopoly." Ph. D. dissertation, Princeton University, 1983.

Akerlof, George A. "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism," *Quarterly Journal of Economics* 84 (August 1970), 488 – 500.

Allen, Franklin. "Capital, Sharecropping, Ability and Information." Mimeographed. Oxford: Oxford University, 1980.

Arnott, Richard, and Joseph E. Stiglitz. "The Welfare Economics of Moral Hazard." Mimeographed. Princeton, N. J.: Princeton University, November 1983 (a).

______, "Moral Hazard and Optimal Commodity Taxation." National Bureau of Economic Research, Working Paper No. 1154, June 1983 (b).

Baumol, William J. *Business Behavior, Value, and Growth*, New York: Harcourt, Brace & World, 1967.

Berle, Adolf A., Jr. "Non-Voting Stock and 'Bankers' Control." *Harvard Law Review* (1926).

______. *Power Without Property*. New York: Harcourt, Brace, 1959.

Berle, Adolf A., Jr., and Gardiner C. Means. *The Modern Corporation and Private Property*. New York: Macmillan, 1932.

Bernheim, B. Douglas, and Michael D. Whinston. "Common Agency." Mimeographed. Stanford, Calif.: Stanford University, 1984.

Blinder, Alan S., and Joseph E. Stiglitz. "Money, Credit Constraints, and Economic Activity." *American Economic Review, Papers and Proceedings* 73 (May 1983), 297 – 302.

Braverman, Avishay, and Joseph E. Stiglitz. "Sharecropping and the Interlinking of Agrarian Markets." *American Economic Review* 72 (September 1982), 695 – 715.

Cary, W. "Corporate Devices Used to Insulate Management from Attack." *Anti-trust Law Journal* (1969 – 70), 318.

"Corporations and Private Property." *Journal of Law and Economics* 26 (June 1983). Special

54 同时，还应该要记住，其他成功的经济体（例如，德国和日本）中的资本控制体制，与美国的资本控制体制的差别很大。值得深入研究的问题是，对这些差异的解释，是否与特定的法律约束或经济的其他方面有关，或者这些差异只是一个历史事件而已。

issue with papers by Thomas G. Moore; George J. Stigler and Clare Friedland; Douglas C. North; R. Hessen; Nathan Rosenberg; Gardiner C. Means; Eugene Fama and Michael C. Jensen; Oliver E. Williamson; Burt Klein; Harold Demsetz; Rita Ricardo-Campbell; Frank Easterbrook and D. Fischel; Leonard W. Weiss; J. S. McGee; and C. Baldwin.

Diamond, Peter A., and Joseph E. Stiglitz. "Increases in Risk and in Risk Aversion." *Journal of Economic Theory* 8 (July 1974), 337 – 60.

Greenwald, Bruce. "Adverse Selection in the Labor Market." *Review of Economic Studies* (1984), in press.

Greenwald, Bruce, and Joseph E. Stiglitz. "Pecuniary and Market Mediated Externalities: Towards a General Theory of the Welfare Economics of Economies with Imperfect Information and Incomplete Markets." National Bureau of Economic Research, Working Paper No. 1304, March 1984.

Greenwald, Bruce, Joseph E. Stiglitz, and Andrew Weiss. "Informational Imperfections in the Capital Market and Macro-Economic Fluctuations." *American Economic Review, Papers and Proceedings* 74 (May 1984), 194 – 200.

Grossman, Sanford J., and Oliver D. Hart. "Takeover Bids, the Free-Rider Problem, and the Theory of the Corporation." *Bell Journal of Economics* 11 (Spring 1980), 42 – 64.

______, "The Allocational Role of Takeover Bids in Situations of Asymmetric Information." *Journal of Finance* 36 (May 1981), 253 – 70.

______, "Corporate Financial Structure and Managerial Incentives." In *Economics of Information and Uncertainty*, edited by John McCall, pp. 107 – 40. Chicago: University of Chicago Press, 1982.

Grossman, Sanford J., and Joseph E. Stiglitz. "Information and Competitive Price Systems." *American Economic Review* 66 (May 1976), 246 – 53.

______, "On Value Maximization and Alternative Objectives of the Firm." *Journal of Finance* 32 (May 1977), 389 – 402.

______, "Stockholder Unanimity in Making Production and Financial Decisions." *Quarterly Journal of Economics* 94 (May 1980), 543 – 66 (a).

______, "On the Impossibility of Informationally Efficient Markets." *American Economic Review* 70 (June 1980), 393 – 408 (b).

Jensen, Michael C., and William H. Meckling. "Theory of the Firm: Managerial Behavior, Agency Costs, and Ownership Structure." *Journal of Financial Economics* 3 (October 1976), 305 – 60.

Keeton, William R. *Equilibrium Credit Rationing*. New York: Garland, 1979.

Larson, John. "Why Long Term Incentives Fail." Paper presented to a Conference on the Economics of Incentive, Cooperation, and Risk Sharing." New York University, March 1984.

Leland, Hayne E., and David H. Pyle. "Informational Asymmetries, Financial Structure, and Financial Intermediation." *Journal of Finance* 32 (May 1977), 371 – 87.

Manne, Henry G. "Mergers and the Market for Corporate Control." *Journal of Political Economy* 73 (April 1965), 110 – 20.

Nalebuff, Barry J., and Joseph E. Stiglitz. "Prizes and Incentives: Towards a General Theory of Compensation and Competition." *Bell Journal of Economics* 14 (Spring 1983), 21 – 43.

Ross, Stephen A. "The Economic Theory of Agency: The Principal's Problem." *American E-*

conomic Review, *Papers and Proceedings* 63 (May 1973), 134 - 39.

______, "The Determination of Financial Structure: The Incentive-Signalling Approach." *Bell Journal of Economics* 8 (Spring 1977), 23 - 40.

Shapiro, C., and Joseph E. Stiglitz. "Equilibrium Unemployment As a Worker Discipline Device." *American Economic Review* 74 (June 1984), 443 - 45.

Stiglitz, Joseph E. "On the Optimality of the Stock Market Allocation of Investment." *Quarterly Journal of Economics* 86 (February 1972), 25 - 60 (a).

______, "Some Aspects of the Pure Theory of Corporate Finance: Bankruptcies and Takeovers." *Bell Journal of Economics and Management Science* 3 (Autumn 1972), 458 - 82 (b).

______, "Incentives and Risk Sharing in Sharecropping." *Review of Economic Studies* 41 (April 1974), 219 - 55.

______, "Information and Economic Analysis." In *Current Economic Problems*, edited by Michael Parkin and A. Robert Nobay, pp. 27 - 52. Cambridge: Cambridge University Press, 1975 (a).

______, "Monopolistic Competition and the Capital Market." IMSSS Technical Report No. 161. Stanford, Calif.: Stanford University, February 1975 (b).

______, "Theory of Local Public Goods." In *The Economics of Public Services*, edited by Martin S. Feldstein and Robert P. Inman, pp. 274 - 333. London: Macmillan, 1977.

______, "Pareto Optimality and Competition." *Journal of Finance* 36 (May 1981), 235 - 51.

______, "Ownership, Control, and Efficient Markets: Some Paradoxes in the Theory of Capital Markets." In *Economic Regulation: Essays in Honor of James R. Nelson*, edited by Kenneth D. Boyer and William C. Shepherd, pp. 311 - 41. Ann Arbor, Mich.: Michigan State University Press, 1982 (a).

______, "Information and Capital Markets." In *Financial Economics: Essays in Honor of Paul Cootner*, edited by William F. Sharpe and Cathryn M. Cootner, pp. 118 - 58. Englewood Cliffs, N. J.: Prentice Hall, 1982 (b).

Stiglitz, Joseph E., and Andrew Weiss. "Credit Rationing in Markets with Imperfect Information." *American Economic Review* 71 (June 1981), 393 - 410.

______, "Incentive Effects of Terminations: Applications to the Credit and Labor Markets." *American Economic Review* 73 (December 1983), 912 - 27.

Tiebout, Charles M. "A Pure Theory of Local Expenditures," *Journal of Political Economy* 64 (August 1956), 416 - 24.

挫败敌手：管理寻租与经济无效*

企业管理理论有一漫长而充满争议的历史，这大概要追溯到伯利（Adolf A. Berle）、米恩斯（Gardiner C. Means）的经典研究（1932）。这类理论假定，职业经理人在运营企业时有相当程度的相机决策权，他们运用这些权力从事其他目标，而非最大化股东的市场价值。

后来的批评者否认经理人实际上拥有相机决策权，认为相机决策行为在均衡时不可能存在。要么办公室内部投票决定管理层，企业被收购；要么在一个竞争性的产品市场上，企业宣布破产。

近来，信息经济学的文献，大大抨击了这类批评。因为职业经理人本身的专业知识和对企业各个方面的了解，他们对公司决策的作用是很重要的。然而，信息的不对称性，同样要求管理层限制董事会加诸于经理人的硬性规定。同时，股东投票机制或者收购机制都不足以提供有效的规则限制[1]。因为不需要付出任何成本就能获得昂贵的信息，从而精明地行使投票权，所以小股东存在利用其他人的努力形成“搭便车”问题。此外，如果收购能增强企业的生产力，则也能同样增加其市场价值。于是，在这种情况下，小股东最好继续持有股票而非出卖股票。所以，当所有权分散时，收购是很困难的。确实，理论的论证揭示，收购机制的实际作用比通常我们想像的效率要低。在解释管理规则与管理决断能力的存在性方面，这也成为一个理

---

* Aaron S. Edlin, Joseph E. Stiglitz, “Discouraging Rivals: Managerial Rent-Seeking and Economic Inefficiencies”, *The American Economic Review*, Vol. 85, No. 5 (Dec. 1995), 1301 - 1312. 本文原本是为 1991 年 5 月经济政策研究中心赞助的一次学术会议而准备。我们感谢国家科学基金会和胡佛研究所的赞助，也感谢 Jeremy Bulow, Eric Emch, Kirsten Landeryou, Peter Klein, Jonathan Paul, Jeffrey Zwiebel，以及匿名审稿人的建议。本文所述观点仅为作者所持，并不一定代表作者所属机构。

1 收购机制方面的理论文献包括 Stiglitz（1972，1975）；（Sanford S. Grossman and Oliver D. Hart）（1980）。实证文献包括 Michael Jensen（1988）；David Scharfstein（1988），Andrei Shleifer and Robert W. Vishny（1988）。对投票控制机制的讨论，请参考 Stiglitz（1981）；Jensen and Richard S. Ruback（1983）；Randall Morck et al.（1989）。

论的困惑。[2]

这些观察为早期的管理文献的视觉寻求支持——这类文献假定经理人拥有管理判断决断能力，问题是他们怎么使用这种能力（James G. March and Simon，1958；Robin L. Marris，1964，Oliver E. Williamson，1964；Harvey Leibenstein，1966）。本文追随那一传统，检验了一个普遍规律——通过诸如投资选择这一手段，管理判断决策能力能为经理人带来取得“租金”的机会（收入超过其机会成本）。我们也由此而扩展了斯莱佛和威斯尼（Shleifer and Vishny，1989）那一富有洞见的文章的主题思想。

先前不完备合约方面的文献强调寻租导致的投资后果。这类文献解释对机会主义的关注，怎样能导致这类“特定关系”投资的投资不足，譬如，经理人对“特定企业”的人力资本的投资（Aaron S. Edlin and Stefan Reichelstein，1993）。我们在这里论证更广视角上的对经理人决策偏见：通常来讲，管理性寻租不仅仅影响投资水平，也影响投资形式。[3]

我们的假说很简单：给定当今公认的管理判断能力，经理人有动力去执行这种能力以增加收入。任何合约都可以重新拟定，经理人的薪酬经常是复杂的管理层——也即董事会——与敌手管理层或者收购者之间讨价还价过程的结果。在这个过程中，两个关键要素是：与竞争对手相比较，现任经理人的生产力，以及竞争对手的要价。职业经理人可以通过增强自己的生产力（这是通常所强调的）或者减低竞争对手的威胁来增加自己的收入。[4] 斯莱佛和威斯尼（1989）考察了后一种可能性，并揭示了职业经理人可以选择对那些他能比对手管理得更好的项目投资，来增加其“租金”。

这里，我们探索了另一条“挖筑壕沟，抵制对手”的路线。斯莱佛和威斯尼（1989）效应最典型的例子是，秘书通过将文件系统安排成仅仅适应自己特定寻找模式这种办法，令自己变得不可替代和“有价值”。然而，我们观察到职业经理人不需这样就能获得工作，即使他们选择了那类不是最

2 早期的文献，例如，伯利（Berle，1962）强调银行的规则纪律性角色（也可参考斯蒂格利茨（1985）），更近的，组织理论的学者和一些经济学家也强调社会控制机制（试图使职业经理人与公司统一认同，并采纳将公司福利作为他们自己私人目标的方法），在这方面，参考阿克洛夫（George A. Akerlof，1991）和西蒙（Simon，1991）。

3 艾德林（Edlin，1992）论证对不完备契约背景下的重要的扭曲，论证扭曲同样也是在于投资形式上的，而非投资水平。

4 尽管文章背景不同，我们在本文的论证与萨勒普（Steven C. Salop）和谢夫曼（David T. Scheffman）（1983，1987）对不完全竞争市场下企业行为的分析很相似。他们争辩说，企业不仅可以通过自己的生产力，也可以通过降低对手的威胁来增强自身的利润（也即是，通过对手的成本）。

适合自己特定能力的投资，他们仍然能够被再次雇用的。我们揭示了，职业经理人通过对信息不对称特别严重的活动进行投资来保住自己的工作（我们这里的论证与 Laurie S. Bagwell and Zechner（1993）的思想相通）。在我们的理论中，经理人设法制造信息不对称，而非利用自己才能。因而，他们是在从事一种比斯莱佛和威斯尼（1989）考虑的那一情形更为消极的某种形式的寻租活动。

我们的理论预测，举例来说，经理人会过度看重承诺某种潜在（但非常确定）协同的（Synergle）并购，因为只要现任经理人能尽早的得到比敌手更准确的协同效应的信号，并购就可以为他带来租金。这一论断暗示，新近的媒体和通信行业的并购浪潮就是一个敏感的“挖筑壕沟，抵制对手”的战略，即使预期的协同效应可能不怎么样，潜在的协同效应是很大的。

这种可能性将随着文章的展开而变得越来越清晰。本文第一部分建立挫败敌手的基本框架；第二部分分析如何挫败那些能观察到投资水平的敌手。第三部分考虑如何挫败那些不能观察到投资组合的敌手；第四部分概括了政策与实证研究分析的某些含义。

一、挫败敌手：框架

我们引入一个三阶段模型。第一阶段，经理人选择投资；第二阶段，董事会决定由谁来管理企业的资产；第三阶段，经理人管理资产。

经理人使他们在第二阶段被再次雇佣的概率与他们被雇佣后期望获得的租金的乘积最大化的投资

$$\max_{\{I_1,I_2\}}[E[\text{租金} \mid \text{再次雇佣}] \times \Pr\{\text{再次雇佣}\}]$$

上面公式并没不意味着风险中性：相反，这些租金是以效用单位计算的，代表了经理人可以从外部机会得到的效用（我们假设是常数）与第二阶段的合约所得到效用之差额，它是经理人所签订合约与随后付出的努力水平的函数。我们马上就会清楚这些租金的来源，但在更宽泛的层面来看，我们的论证很简单：在其他条件不变的情况下，随着经理人的竞争敌手的内在能力的下降，或者他们努力水平下降，或者他们对接受工作职位要求的薪水上升，现任的经理人将能从合约的谈判中获得更多的租金。

当然，一些研究公司文化的学生对董事会与经理人谈判的这一现象抱有不同意见：经理人通常被认为是合约条款的制定者，董事会只是象征性的批准通过。的确，董事会有时候批准的与现任经理人的合约只会让公司的管理

更差。不过，即使那些对董事会能否代表股东利益持极端意见的人也不得不承认，某些条件确实限制了经理人的薪金索求。我们指出，这些限制最终来自于被竞争敌手取代的威胁。因此，我们研究经理人如何通过挫败敌手来减少这些威胁。

在我们的表述中，企业的期终或者说第三阶段的回报 R 是经理人的有效质量与企业在建项目的前景的乘积：

$$R = (C_1 + C_2)q(e) \tag{1}$$

这里，q 是经理人的有效质量，这可能依赖于努力水平 e，C_i 代表项目 i 的前景，$i \in \{1, 2\}$，两个项目可以看为一个“短期项目”，另一个为“长期项目”；或者作为同一时间进行的两种不同活动。C_i 也可能代表潜在的现金流或项目 i 的基础“资本”。它以风险 θ_i 随机的联系到企业第一阶段的投资，并随着投资的增加按比例地增加：

$$C_i = F_i(I_i) + \theta_i \tag{2}$$

其中，$\theta_i \sim N(0, I_i^2\sigma_i^2)$，$i \in \{1, 2\}$

我们考虑两种基本的情形：最初，外部人可以观察到两种投资活动的组合；然后，他们只能观察到总投资 $I_1 + I_2$。在两种情况下，现任经理人都能准确地观察到 C_i，而他的竞争敌手必须从信号 Y_i 来推断 C_i，我们有如下定义

$$Y_i = C_i + \varepsilon_i \tag{3}$$

其中，$\varepsilon_i \sim N(0, I_i^2 w_i^2)$，$i \in \{1, 2\}$ 同时 ε 和 θ 是相互独立的。

投资的最优组合程度最大化 $E[C_1 + C_2]$。假设有内点解，最优投资是：有 I_1^* 和 I_2^* 使得 $F_1'(I_1^*) = F_2'(I_2^*)$ 成立。然而，下面的两部分将揭示，当经理人试图为自己设置保护措施时，最优的投资并不能解决投资问题。不管经理人选择的投资组合对外部人来说是可观察的或者不可观察的，投资选择都是扭曲的。这两种情形下，尽管扭曲的理由不尽相同，但是经理人比较偏好那种对他自己比外部人能更快或更全面掌握情况的投资。

二、情形1：可观察的投资

竞争敌手以他们对观察到变量 Y_1 和 Y_2 为条件，形成他们对企业的潜在现金流的期望。这一期望由下式给定：

$$E[C_1 + C_2 \mid Y_1, Y_2] = \beta_0 + \beta_1 Y_1 + \beta_2 Y_2 \tag{4}$$

其中，$\beta_i = \dfrac{\sigma_i^2}{w_i^2 + \sigma_i^2}$ $i \in \{1, 2\}$，$\beta_0 = \dfrac{w_1^2}{w_1^2 + \sigma_1^2}F_1(I_1) + \dfrac{w_2^2}{w_2^2 + \sigma_2^2}F_2(I_2)$

我们将在附录里推导。因为在这一部分的分析中，外部人可以观察到投资组合，插入项 β_0 是时间投资 I_1 和 I_2 的函数。这一特色与第三部分分析不可观察到的投资情况不同。

1. 努力挫败敌手的投资

选择好投资后，资产必须有人来管理，或者由现任经理人，或者由他的敌手。在这两种情况下，给定经理人的信息，他选择让预望效用函数最大化的努力水平。企业必须提供给经理人某些激励以促使他们付出努力，为简单起见，我们假设负责管理资产的经理人被授予一份分享利润的激励合约。[5]因此，如果 W 代表经理人 j 的固定工资收入，s 代表他的利润份额，U 代表他的效用函数，p_je 是他付出努力水平 e（p_j 是他能力的一个指标）的个人成本，而 Φ 是他的信息集。经理人选择努力水平以最大化

$$E[U[W+s(C_1+C_2)q(e)-p_je] \mid \Phi] \tag{5}$$

假定绝对风险厌恶（constant absolute risk aversion）系数为常数，这使得确定性等价（certainty equivalent）对于财富的均值与方差是线性的，[6] 经理人的问题是最大化

$$W+sq(e)E[(C_1+C_2) \mid \Phi]-ks^2q^2(e)\mathrm{Var}[(C_1+C_2) \mid \Phi]-p_je \tag{6}$$

这里 $\mathrm{Var}[(C_1+C_2) \mid \Phi]$ 是 C_1+C_2 对经理人信息集 Φ 的条件方差，k 是绝对风险厌恶程度的一半。我们假设有效的工人质量 $q(e)$ 随努力水平而增加。努力是有成本的，因此，在其他条件等同的情况下，更高努力成本 p_j 的经理人将发挥更少的努力，而且有效质量也更低。然而，经理人的努力，以及他的有效质量，同样依赖于他对企业前景的评估。

在我们模型中，当敌手接管了之后，他必须在知道前景 C_1+C_2 之前，决定自己的努力水平。这一时间上的规定抓住了一个重要特色，否则的话，这一特色就会由于我们假设了敌手选择努力之前可以观察到 C_1 与 C_2 而被忽略掉。当敌手接管时，他不会将包含信息“C_1”与“C_2”的关键交到某些安全之处。这是对私有信息的一大讽刺。注意到斯考利（John Sculley）花了4个月才发现频段的条件，并且，要考虑这类信息通常比那类情形中还要来得更为抽象。[7] 相关信息可能是相当分散的，包括很多前

5 对线性激励计划的一个支持论证，参考 Bengt Holmstrom and Paul R. Milgrom（1987）。

6 我们没有要求所有的常绝对风险规避结构以使引理1成立和随后的分析有效。特别地，强线性并不重要：从我们的目的来说，努力水平分开地进入经理人的效用函数，以及经理人的效用函数是期望财富的增函数，是财富方差的减函数，这就足够了。

7 见 Jonathan Weber（1994）对斯考利频谱传奇故事的描述。

任或者当前的企业主管，而新管理层对这些信息的了解远远不能够做到及时。因此，当我们极端的假设——经理人取得管理职位之后，就没有什么新信息可以得到——他也很简洁的抓住了这一事实：对于前任所知道的大部分信息新管理者是不可能知道的。我们现在准备陈述一个有用的引理：

引理1：努力的成本和其他要素不变，经理人 j 的努力选择问题 e^{*j} 的一个内点解随 $E[(C_1+C_2)\mid\Phi]$ 递增，并随 $\mathrm{Var}[(C_1+C_2)\mid\Phi]$ 递减。

证明：引理可以从经理人的目标函数的形式，以及有效质量 q 随着努力水平递增这一事实推导出来。$E[(C_1+C_2)\mid\Phi]$ 对 q 的偏导数为正，而 $\mathrm{Var}[(C_1+C_2)\mid\Phi]$ 对 q 的偏导数为负，于是米尔格拉姆和香农（Milgrom and Shannon, 1994）的单调选择定理隐含着：经理人 j 的努力选择问题的解对于 $E[(C_1+C_2)\mid\Phi]$ 是弱增的，对于 $\mathrm{Var}[(C_1+C_2)\mid\Phi]$ 是弱减的。这个结果是很严格的，只要 e^{*j} 是假定内点的，因为一旦一阶条件对某一期望或者方差成立，它们对其他更高的值就不再成立。

我们接下来验证，通过降低敌手管理的有效性，当前的管理层的行动怎样可以损害应聘者对经理人职位的挑战。现任的经理人可以通过增加敌手对企业前景的不确定性而降低敌手的努力。相对于敌手，现任者的努力以及有效质量随着下式的增加而上升

$$\mathrm{Var}[(C_1+C_2)\mid\Phi_{riv}]-\mathrm{Var}[(C_1+C_2)\mid\Phi_{inc}] \qquad (7)$$

这里 Φ_{riv} 表示敌手的信息而 Φ_{inc} 则表示现任经理人的信息。

记住现任经理人能观察到 C_1+C_2，于是 $\mathrm{Var}[(C_1+C_2)\mid\Phi_{inc}]=0$，对比起来，敌手的管理团队就只能从噪音信号 Y_1 和 Y_2 来推断 C_1+C_2，在本文附录，我们推导了下式

$$\mathrm{Var}[C_1+C_2\mid Y_1,Y_2]=\frac{\sigma_1^{\ 2}w_1^{\ 2}I_1^{\ 2}}{\sigma_1^2+w_1^2}+\frac{\sigma_2^{\ 2}w_2^{\ 2}I_2^{\ 2}}{\sigma_2^{\ 2}+w_2^{\ 2}} \qquad (8)$$

式（8）揭示了经理人将喜欢选择噪音更混杂的投资（对任何给定的 σ^2，w^2 将更高）：这样的投资增加了式（7），而减少了外部竞争对手的有效性[8]。从增加的噪音信号获得的收益随着基本奉献的增高而增大，在

8 注意，这一结果与标准的 Jensen and William H. Meckling（1976，第4.1部分）的资产替代结果是相当不同的。他们论证，股票持有者对投资到高风险的资产是有偏见的，因为防止不利趋势的能力有限。噪音 w^2 应该从基本风险 σ^2 中区分开来。

$\sigma^2=0$的极端的情形，因为潜在现金流完全由可观察到的投资水平决定，w^2就变得无关紧要了。最后，由于 I^2 的负责性，在选择风险（σ^2）以及对于外部者的噪音信号为（w^2）的相似的项目时，现任经理人可以通过专注于他们本身的投资来降低敌手的有效性。相应地，寻租行为就可能会导致过度专业化。

2. 挫败应聘者

除了降低敌手的有效质量，现任经理人的投资选择也可以提高敌手的薪金叫价。对于保留效用水平为 r（j）的经理人 j，要他愿意接受为企业工作，起码要提供足够高的如下薪金

$$W + sq(e^{*j})E[(C_1 + C_2) \mid \Phi] - p_j e^{*j} > r(j) + ks^2q^2\mathrm{Var}[(C_1 + C_2 \mid \Phi] \tag{9}$$

上式的右边代表一个大家熟悉的事实：必须补偿经理人以保留效用 r（j），以及激励计划的风险由 s 表示。当 Var（$C_1+C_2 \mid \Phi_{riv}$）增加时，任何给定的补偿方案将降到低于某些敌手的保留工资水平以下。这些作为竞争对手的经理人于是选择不应聘，进而减弱了竞争对手的总体质量并增加了现任经理人被再次雇佣的机会。

3. 增加敌手对“赢者的诅咒”

经理人投资于更混杂的噪音项目的第三个理由是，这样的投资项目制造了“赢者的诅咒”效应。这一效应进一步挫伤了应聘者应征的热情，并迫使那些真正的应聘者索取一个更高的“不熟悉环境”的补偿。记住在位者知道C_1+C_2，于是

$$E(C_1 + C_2 \mid \Phi_{inc}) = C_1 + C_2 \tag{10}$$

他的敌手不知道 C_1+C_2，只知道如果他准备观察 Y_1 和 Y_2 的值，以使得他自己的期望与在位者一样的话，董事会将更可能保留现任者，我们上面讨论过这么做的两个理由：（a）其他情况相同，现任经理人工作更加努力，因为对他们付出的回报比较确定；（b）现任经理人愿意接受更低报酬的工作，因为他们可感知的风险更小。

因此，现任经理人能够与董事会达成一个协议，打败敌手对管理职位的挑战，除非：（a）现任经理人比起敌手来，保留工资太高；（b）董事会对敌手能力的评估比现任经理人高；（c）敌手过度评估了项目的价值（ε 的值很高）。保持前两个要素不变，敌手知道只有当他过度乐观时，才能得到管理职位，他把这一情况在形成期望时一并加以考虑（考虑《华尔街日报》在阿克斯（John Akers）安静离开后的头条新闻：大标题

是“阿克斯离开 IBM”，小标题“谁想接任？”（Michael W. Miller and Laurence Hooper，1993）。[9]

在决定是否要应聘时，敌手事实上考虑的是

$$E[C_1 + C_2 \mid Y_1, Y_2, \text{不再次雇佣}] \tag{11}$$

而非

$$E[C_1 + C_2 \mid Y_1, Y_2] \tag{12}$$

这一对“赢者的诅咒”，进一步降低了敌手的努力和他们接受职位的意愿，[10]就像我们上面解释的

$$E[C_1 + C_2 \mid Y_1, Y_2, \text{不再次雇佣}] < E[C_1 + C_2 \mid Y_1, Y_2] \tag{13}$$

假定，现任经理人以被再次雇佣为前提条件的租金线性随着楔子增加，则

$$C_1 + C_2 - E[C_1 + C_2 \mid Y_1, Y_2, \text{不再次雇佣}] \tag{14}$$

以及条件方差

$$\text{Var}[C_1 + C_2 \mid Y_1, Y_2, \text{不再次雇佣}] \tag{15}$$

记住现任经理人的投资问题是

$$\max_{\{I_1, I_2\}} [E[\text{租金} \mid \text{不再次雇佣}] \times \Pr\{\text{不再次雇佣}\}] \tag{16}$$

我们的分析指出，对与噪音投资的回报（即是给定任何 σ^2 具有更高 w^2 值的情形）要大于那些 Y 能够起到很好潜在现金流信号作用的项目的回报，这有三个不同的原因：（a）噪音信号直接挫败了敌手的努力，并因此使得敌手成为经理人的吸引力更小；（b）噪音直接挫败了敌手应聘的意愿（并迫使他们索取高的薪金），从而使得他们不可能成为可行的威胁；（c）噪音增加了信息的不对称性以及相应的“赢者的诅咒”效应。

考虑一下如下极端的情形：敌手能够完全观察到项目 1，且只能通过噪音信号来观察项目 2。如果两个项目的基本面是充分接近的，我们将预计项目 2 被选择，即使股东可能比较喜欢投资在项目 1。通过投资在噪音混杂的项目，经理人提高了被再次雇佣的概率与被再次雇佣时他们可以获得的平均租金。第四部分论证了近来的某些为新出现的信息时代而做准备的并购，可能预示着这样的“挖沟堑壕”战略活动。

---

9 Miller and Hooper（1993）指出：“最终，IBM 公司痛苦的弄清楚了，完全没有必要搞一次难以驾驭的董事会政变。”

10 敌手的管理团队对相关方差的评估，在考虑“赢者的诅咒”这一情况是不得不做些变更，但是我们假定，切断这一可能将使按风险排序的投资保持不变。

4. 扩展：多维努力水平

我们可以做类似的分析来探索另一个模型的情况：与在第一阶段选择投资的情况不同，经理人选择如何在两个项目之间配置自己的努力。保罗(1992) 考察了以股票为基础的激励补偿效率的情况下的这一问题。根据我们的符号表达方式，他发现经理人的努力将随 σ_i^2 的递增，而随 w_i^2 递减。而我们对于这一分析的结论却与他的结论有所不同，因为与他不同的是，我们假设管理合约在这些决策之后是可以再拟定的。噪音投资（高的 w_i^2）可能将挫败敌手而同时将现任经理人摆到更加有利的位置上，从而能在再拟定合约的过程中抽取更多的租金。比较起来，如果合约是事先设定以包含股票的补偿方案，并且不能再拟定的话，保罗的效应将占主导地位，因为就像保罗所观察到的，噪音投资项目有某项缺点：它们对股票的价格几乎没有什么影响。

三、情形 2：不可观察到的投资

我们假设，经理人有一预算 $\bar{I}$ 可以分配到项目 1 和项目 2，而且，尽管财务报表显示项目总投资，它却没有披露两个项目的组合比例投资额度。因此，敌手对企业的前景的评估必须综合考虑可观察的信号 Y_1 和 Y_2，以及均衡的投资比例来决定。我们在附录里推导了以观察值 Y_1 和 Y_2 为条件的前景的期望值，现列出如下

$$E[C_1 + C_2 \mid Y_1, Y_2] = \beta_0 + \beta_1 Y_1 + \beta_2 Y_2 \tag{17}$$

其中，

$$\beta_i = \sigma_i^2/(\sigma_i^2 + w_i^2) \quad i \in \{1,2\}, \beta_0 = \frac{w_1^2}{w_1^2 + \sigma_1^2}F_1(\hat{I}_1) + \frac{w_2^2}{w_2^2 + \sigma_2^2}F_2(\hat{I}_2)$$

改变投资组合将同时改变 Y_1 和 Y_2，但由于敌手不能观察到这些变动，它们将不会改变式（17）给出的形成期望的规则。这里的插入项是均衡投资水平 $\hat{I}_1$ 和 $\hat{I}_2$ 的一个函数，这跟式（4）有所不同，因为那里的情况是实际投资是可以观察到的。

通过提高 $E[C_1 + C_2 \mid \Phi_{inc}] - E[C_1 + C_2 \mid \Phi_{riv}]$，现任经理人可以减低敌手的有效性。我们可以看到，通常他不会选择最优的投资水平 I_1^* 和 I_2^*，因为如果他将更多预算投入到更为混杂的噪音项目（即是更高的 $\frac{w^2}{\sigma^2}$ 比值和更低的 β 系数）中去，他将获益更多。而这一转移增加了楔子 $C_1 + C_2$

$-E[C_1+C_2 \mid Y_1, Y_2]$：[11]要看清这一点，观察到，作为从 I_1^* 和 I_2^* 投资转移到更混杂的噪音项目一微小变差的一阶渐近，对于任何给定的 ε_i 和 θ_i，当 C_1+C_2 保持不变时，因为 $F_1'(I_1^*)=F_2'(I_2^*)$，$E[C_1+C_2 \mid Y_1, Y_2]$ 随之下降。提高投资到更混杂的噪音项目的水平，将降低敌手最优努力水平，从而降低他们的有效质量。"赢者的诅咒"问题类型于前一节，加剧恶化了这一效应。事实上，在混杂的噪音项目里面，所有的投资水平都可能是潜在的均衡。然而，充分递减的真实回报这一特性，最终，当噪音项目的资产回报更低时，经理人将不再投资以制造噪音，提高更低的投资 C_1+C_2 到一阶渐近水平，就好像它降低了 $E[C_1+C_2 \mid Y_1, Y_2]$。给定任何激励契约，降低 C_1+C_2，也降低了现任经理人的努力，这使得他成为一个不太合意的经理人。此外，某种意义上，当经理人在决策时看重 C_1+C_2 这一要素时，降低 C_1+C_2 是非常不合意的。

最后，我们应该注意到，由于这里敌手不能观察到投资组合，现任经理人不能影响 $\mathrm{Var}[C_1+C_2 \mid Y_1, Y_2]$，而在上一部分，他则可以这么做。在这一部分，$\mathrm{Var}[C_1+C_2 \mid Y_1, Y_2]$ 是由均衡投资水平决定的。

四、含义

我们以一些解释和含义来总结我们的分析，并提出如下四个问题：

（1）这一模型是否解释了管理者对过度增长的癖好，就像唐纳森（Gordon Donaldson，1984，pp. 36－42）观察到的那样？

在我们看来，经理人的租金是与噪音项目的投资相联系的。如果一家企业从来不投资于新项目而只是保持旧资产，它的经理人只能取得很少的租金，因为过去的项目将能为资产的价值提供很好的信息。

然而，如果不对增长加以区分，这也不能帮助经理人。例如，重复已建的确知价值的项目并不能让经理人受益。在我们模型中，当回报是不确定的，而且对这种不确定性的了解，内部人比外部人快的话，增长就能为经理

---

11 现任经理人形成关于外部人的未知悉的期望的预期，将是使用他对投资水平的完全信息。现任经理人的期望可以写成 $E_{in}[E_{out}(C_1+C_2 \mid Y_1, Y_2) \mid I_1, I_2]$，要认识到在均衡的投资水平 $\hat{I}_1$ 和 $\hat{I}_2$ 下，均衡条件

$E_{in}[E_{out}(C_1+C_2 \mid Y_1, Y_2) \mid \hat{I}_1, \hat{I}_2]=E_{in}[(C_1+C_2) \mid \hat{I}_1, \hat{I}_2]$，这是很重要的。如果现任经理人再配置投资来试图制造楔子以获取租金时，因为两个期望算子的信息是不同的，这一等式于非均衡情况将不再成立。这一等式是反映一致性或者理性预期的条件。

人增加租金收入。我们认为，比较普遍的情况是，在一些新兴产业，σ^2 和 w^2 都相对较高，所以，我们应该对这些产业投资寻找“挖沟堑壕”战略活动的证据。补充了更多传统的“实证建设”之后，模型因此也能更好地解释经理人对增长的癖好。

对创新活动的癖好，当然也不是全然不好的。对创新活动的癖好以及在 R & D 活动上的支出，有助于平衡过度的保守倾向，以及诸如“管理近视症”，眼光短浅之类的倾向，[12]也有助于平衡经理人通常比“风险多样化分散”后的股东更趋于厌恶风险这一事实。

（2）就公司收购与合并的效率而言，这一理论又有什么含义呢？

传统的收购理论论证说，20 世纪 80 年代典型的公司重组必定是增进效率的，因为在每种情况下，资产被卖到能最好管理这些资产的经理人或管理团队手上，这一点可以由中标者愿意付出的价格证明。我们对信息不对称的分析，丰富了对并购模式的理解，也解释了低效率是如何产生的。如果“成熟”的企业意味着对经理人来说租金的稀缺的话，则企业的资产所有者（他要最大化价值）就不得不自己接手管理它。非资产所有者的经理人可以用收益去收购那些他的敌手难以估计价值的资产。因此，在资产拍卖中（譬如，最近以来 QVC 和 Viacom 对 Paramount 的收购提议），企业愿意支出多少，不仅依赖于企业能够从资产中获得多少，也依赖于经理人自己能从控制这些资产中得到多少租金。如果经理人能够快速地将资产的市场价值变得对外部人来说更不确定的话，那么租金也将更多。因而，资产的“管理”价值也将因企业而异，因此，资产很可能被卖到管理价值最高的企业，而不一定是卖给期望回报值最高的企业。对企业已存资产能够带来巨大潜在协同效应的并购，带来的管理性租金可能是相当高的。由于回报是随着协同而实现的，不可能通过研究目标企业的历史记录而确定。只有时间才能够揭示回报，此外，现任经理人对回报的掌握情况要比外部敌手更快。

新近的电信、计算机、娱乐产业的并购浪潮很可能促进“挖沟堑壕”战略。[13]经理人希望，21 个世纪的关键信息技术就是他们正在寻找的长期协

---

12 对近视症的简短评述，请参考 Paul（1991），他的文章对为何当处于混杂环境时，经理人集中关注于短期回报以最大化股票价值，提供了一个指导性的理论。

13 例子包括时代—华纳的合并，Viacom 和 Paramount 的合并，以及 AT&T 和 McCaw Cellular Communications 的合并。参考 1994 年《华尔街日报》对新一轮并购浪潮的报道。

同的一个机会。这类并购有巨大的上升潜力，但对管理层来说，减缩趋势也同样重要。正是“潜在性”的内在不确定性创造了租金，只要这一“潜在性”保持不变，再加上现任经理人比外部敌手能更早了解到实际情况。现任经理人就能预期到个人的优势和收益。对管理者来说，购买这类资产是明智的并购，即使它的期望回报普普通通。比较起来，非协同式的合并，譬如电信的20世纪60、70年代的多样化资产积聚的形成，从“挖沟堑壕”的意义上来说，价值并不高；它们导致了80年代某种形式的收购与企业资产所有权人自决的主导地位。

(3) 对于我们讨论的寻租行为，有什么方法克服？

首先，注意到管理行为中的自然偏差，董事会或者股东希望能发现并阻止那些最明目张胆的有偏差的投资行为。最近的一个例子就是，QVC公司的两个大股东Liberty Media 和 Comcast 挫败了迪勒（Barry Diller）对CBS的并购图谋（*Chicago Tribune*，1994）。然而，企业面临的一大困惑是，能够增加利润的投资，同时也增加了管理者的租金。新兴企业的利润可能挺高，不过，那些有助于企业提高利润壁垒的同时也增加了管理者的租金（对于不同的见解，参见Benjamin E. Hermalin，1992）。

其次，董事会也可能寻求一种限制薪酬再拟定的机会，它也可能在公司的规章制度中将CEO的薪酬制度化。然而，这种方法也有缺点：规章制度的惯性力量本身也是有成本的（美国烟草公司的例子就很好地揭示了这种缺乏弹性的风险）。[14]美国某些州已经立法限制公司章程里面的薪酬方案。[15]可能限制（尽管不能消除）寻租的更好的办法，就是对付其他代理问题的普通方法：以低效率的风险投资分担为代价，董事会强迫管理者接受更多的股份。

第三种办法是，由董事会或者由政府设定更为严格细致的核算与报告的硬性规定。譬如，如果报告中清晰地辨别了投资的组合情况，那么第三部分所分析的扭曲就消除了。此外，就某种程度上来讲，如果经理人被强制要求

14 1912年，美国烟草公司采纳了一公司章程，规定公司总裁可以分享净利润的2.5%作为奖金，而每个副总裁则可以分到1.5%。18年后的1930年，这一规则带给总裁的奖金达到84万美元之多。少数股东可能（“Rogers v. Hill” p. 289，U. S. 582，1933）指控他们的不光彩做法实际上是浪费金钱。

15 考虑北卡罗来纳州的商业公司法（North Carolina Business Corporation Act，§16（3），1973）：以公司收入或者商务数量计算的公司主管薪酬方案在采纳5年之后才能生效，除非得到公司发行股票——不管这些股票是否具有投票权——持有者的大多数的投票支持，才能生效。

更快、更准确地报告内部消息的话，那么扭曲也能降低。

我们认识到，正是信息私有的性质使得信息的披露要比全部呈报更为容易。不过，报告的强制要求有一效应。很少有人质疑，正是由于《1933 年美国证券法案》和《1934 年证券交易法案》，才使得重要的信息能为公众所掌握，否则的话，它只为私人拥有（比较美国上市交易公司的信息的可获得性与那些没有这类立法，或者法案没有被执行的国家）。

（4）最后，我们的研究为哪些实证问题打开了方便之门？

我们的文章提出了一连串的问题，而这些问题都需要聪明的研究者来回答。原则上，如果你愿意站在——哪种投资的回报对现任经理人来说比外部敌手了解得更清楚——这一立场上，那你可以验证模型的预测。你也可以尝试去确定是否这类投资的边际回报比那些对内部人和外部人来说一样可以观察到的投资的边际回报要更低。这一结果意味着，经理人偏好混杂的噪音投资，并且，这些投资的水平超过了均衡的回报的分界线。由于其他的文章（比如说，Paul，1991，1992）预测，那些为最大化股东价值服务的经理人，将不敢从事混杂的噪音投资项目，实证结果将有助于理解经理人的目标。

研究投资模式对由财务会计准则委员会（FASB）或者诸如证券交易所（SEC）等管理机构制定报告准则的变化引起的反应同样是有价值的。事实上，巨大的反应可能就意味着经理人实际上盘剥了信息不对称所带来的租金。

附录：

我们先证明文章中推导条件期望 $E[C_1+C_2 \mid Y_1, Y_2]$ 和方差 $\mathrm{Var}[C_1+C_2 \mid Y_1, Y_2]$的一个必要的统计引理。

引理 A1：如果 x 和 y 是独立的正态变量，并且 $x \sim N(0, \sigma^2)$，$y \sim N(0, w^2)$，则给定 $x+y=s$，x 的分布是

$$N\left(s\frac{\sigma^2}{\sigma^2+w^2}, \frac{\sigma^2 w^2}{\sigma^2+w^2}\right)$$

证明：x 的条件密度可以写成如下

$$f(x \mid x+y=s)=\frac{f(x, x+y=s)}{f(x+y)}=\frac{\frac{1}{2\pi\sigma w}\exp\left(-\frac{x^2}{2\sigma^2}\right)\exp\left(-\frac{(s-x)^2}{2w^2}\right)}{\frac{1}{\sqrt{2\pi}\sqrt{\sigma^2+w^2}}\exp\left(\frac{-s^2}{2(\sigma^2+w^2)}\right)}$$

$$= \frac{\sqrt{\sigma^2 + w^2}}{\sqrt{2\pi}\sigma w}\exp\left[\frac{-x^2(\sigma^2 + w^2)}{2\sigma^2 w^2} + \frac{xs}{w^2} - \frac{s^2}{2w^2} + \frac{s^2}{2(\sigma^2 + w^2)}\right]$$

$$= \frac{\sqrt{\sigma^2 + w^2}}{\sqrt{2\pi}\sigma w}\exp\left[-\frac{\sigma^2 + w^2}{2\sigma^2 w^2} \times \left(x - s\frac{\sigma^2}{\sigma^2 + w^2}\right)^2\right]$$

因此，给定 $x + y = s$，x 的分布是

$$N\left(s\frac{\sigma^2}{\sigma^2 + w^2}, \frac{\sigma^2 w^2}{\sigma^2 + w^2}\right)$$

我们可以使用引理 A1 来推出 $E[C_1 + C_2 \mid Y_1, Y_2]$ 和 $\mathrm{Var}[C_1 + C_2 \mid Y_1, Y_2]$，因为已知投资水平

$$\begin{aligned} E[C_1 + C_2 \mid Y_1, Y_2] &= F_1(I_1) \\ &+ F_2(I_2) + E[\theta_1 \mid \varepsilon_1 + \theta_1 = Y_1 - F_1(I_1)] \\ &+ E[\theta_2 \mid \varepsilon_2 + \theta_2 = Y_2 - F_2(I_2)] \end{aligned}$$

根据引理 A1，这等于

$F_1(I_1) + F_2(I_2) + \beta_1(Y_1 - F_1(I_1)) + \beta_2(Y_2 - F_2(I_2))$，其中，$\beta_i = \frac{\sigma_i^2 I_i^2}{(\sigma_i^2 + w_i^2) I_i^2} = \frac{\sigma_i^2}{\sigma_i^2 + w_i^2}$，选择 $\beta_0 = F_1(I_1)[1 - \beta_1] + F_2(I_2)[1 - \beta_2]$，我们就得到文中 $E[C_1 + C_2 \mid Y_1, Y_2]$ 的公式，同样我们有

$$\begin{aligned} \mathrm{Var}(C_1 + C_2 \mid Y_1, Y_2) &= \mathrm{Var}[\theta_1 \mid \varepsilon_1 + \theta_1 = Y_1 - F_1(I_1)] \\ &+ \mathrm{Var}[\theta_2 \mid \varepsilon_2 + \theta_2 = Y_2 - F_2(I_2)] \end{aligned}$$

根据引理 A1，这等于 $I_1^2\sigma_1^2$ 与 $I_1^2 w_1^2$ 以及 $I_2^2\sigma_2^2$ 与 $I_2^2 w_2^2$ 的平均值的一半，就像文中

$$\frac{I_1^2\sigma_1^2 w_1^2}{\sigma_1^2 + w_1^2} + \frac{I_2^2\sigma_2^2 w_2^2}{\sigma_2^2 + w_2^2}$$ 所表达的。

参考文献

Akerlof, George A, "Procrastination and Obedience." *American Economic Review*, May 1991 (*Papers and Proceedings*), 81 (2), pp. 1 – 19.

Bagwell, Laurie S. and Zechner, Josef, "Influence Costs and Capital Structure." *Journal of Finance*, July 1993, 48 (3), pp. 975 – 1008.

Berle, Adolf A., Jr. "Non-voting Stock and Bankers' Control." *Harvard Law Review*, April 1926, 39 (6), pp. 673 – 93.

Berle, Adolf A., Jr. and Means, Gardiner C. *The modern corporation and private property*. New York: Macmillan, 1932.

Chicago Tribune. "QVC Accepts Sweetened Takeovers." Business Section, 5 August 1994,

p. 3.

Donaldson, Gordon. *Managing corporate wealth: The operation of a comprehensive financial goals system.* New York: Praeger, 1984.

Edlin, Aaron S. "Specific Investments, Holdups and the Efficiency of Contract Remedies." Unpublished manuscript, Stanford University, 1992.

Edlin, Aaron S. and Reichelstein, Stefan. "Holdups, Standard Breach Remedies, and Optimal Investment." IBER Working Paper, University of California, Berkeley, 1993; *American Economic Review* (forthcoming).

Grossman, Sanford J. and Hart, Oliver D. "Takeover Bids, the Free-Rider Problem, and the Theory of the Corporation." *Bell Journal of Economics*, Spring 1980, 11 (1), pp. 42 – 64.

Hermalin, Benjamin E. "The Effects of Competition on Executive Behavior." *Rand Journal of Economics*, Autumn 1992, 23 (3), pp. 360 – 65.

Holmstrom, Bengt and Milgrom, Paul R. "Aggregation and Linearity in the Provision of Intertemporal Incentives." *Econometrica*, March 1987, 55 (2), pp. 303 – 28.

Jensen, Michael C. "Takeovers: Their Causes and Consequences." *Journal of Economic Perspectives*, Winter 1988, 2 (1), pp. 21 – 48.

Jensen, Michael C. and Meckling, William H. "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure." *Journal of Financial Economics*, October 1976, 3 (4), pp. 305 – 60.

Jensen, Michael C. and Ruback, Richard S. "The Market for Corporate Control: The Scientific Evidence." *Journal of Financial Economics*, April 1983, 11 (1 – 4), pp. 5 – 50.

Leibenstein, Harvey. "Allocative Efficiency vs. 'X-Efficiency'." *American Economic Review*, June 1966, 56 (3), pp. 392 – 415.

March, James G. and Simon, Herbert A. *Organizations.* New York: Wiley, 1958.

Marris, Robin L. *The economic theory of managerial capitalism.* Glencoe, IL: Free Press, 1964.

Milgrom, Paul R. and Shannon, Chris. "Monotone Comparative Statics." *Econometrica*, January 1994, 62 (1), pp. 157 – 80.

Miller, Michael W. and Hooper, Laurence. "Akers Quits at IBM Under Heavy Pressure; Dividend Is Slashed. Outsiders Will Lead Search for New Chief Executive To Be a 'Change-Master.' But Who'd Want This Job?" *Wall Street Journal*, 27 January 1993, p. A1.

Morck, Randall; Shleifer, Andrei and Vishny, Robert W. "Alternative Mechanisms for Corporate Control." *American Economic Review*, September 1989, 79 (4), pp. 842 – 52.

Paul, Jonathan M. "Managerial Myopia and the Observability of Future Cash Flows." Unpublished manuscript, University of Michigan, 1991.

______, "On the Efficiency of Stock-Based Compensation." *Review of Financial Studies*, 1992, 5 (3), pp. 471 – 502.

Salop, Steven C. and Scheffman, David T. "Raising Rivals' Costs." *American Economic Review*, May 1983 (*Papers and Proceedings*), 73 (2), pp. 267 – 71.

______, "Cost-Raising Strategies." *Journal of Industrial Economics*, September 1987, 36 (1), pp. 19 – 34.

Scharfstein, David. "The Disciplinary Role of Takeovers." *Review of Economic Studies*, April 1988, 55 (2), pp. 185 – 200.

Shleifer, Andrei and Vishny, Robert W. "Value Maximization and the Acquisition Process." *Journal of Economic Perspectives*, Winter 1988, 2 (1), pp. 7 – 20.

______, "Management Entrenchment: The Case of Manager-Specific Investments." *Journal of Financial Economics*, November 1989, 25 (1), pp. 123 – 39.

Simon, Herbert A. "Organizations and Markets." *Journal of Economic Perspectives*, Spring 1991, 5 (2), pp. 25 – 44.

Stiglitz, Joseph E. "Some Aspects of the Pure Theory of Corporate Finance: Bankruptcies and Takeovers." *Bell Journal of Economics*, Autumn 1972, 3 (2), pp. 458 – 82.

______, "Information and Economic Analysis," in Michael Parkin and Avelino R. Nobay, eds., *Current economic problems*. Cambridge: Cambridge University Press, 1975, pp. 27 – 52.

______, "Ownership, Control, and Efficient Markets: Some Paradoxes in the Theory of Capital Markets," in Kenneth D. Boyer and William G. Shepherd, eds., *Economic regulation: Essays in honor of James R. Nelson*. East Lansing, MI: Institute of Public Utilities, Michigan State University, 1981, pp. 311 – 41.

________, "Credit Markets and the Control of Capital," *Journal of Money, Banking, and Credit*, May 1985, 17 (1), pp. 133 – 52.

Wall Street Journal, "Big is Back in Style as Corporate America Deals, Buys, and Merges," Western Edition, 4 August 1994, p. A1.

Weber, Jonathan, "Saga of Apple's Ex-Chief Takes an Unusual Turn." *Los Angeles Times*, Home Edition, 8 February 1994, p. A1.

Williamson, Oliver E. *The economics of discretionary behavior: Managerial objectives in a theory of the firm*. Englewood Cliffs, NJ: Prentice-Hall, 1964.

工资与价格分布理论

均衡工资分布*

长期以来，经济学都用不完美信息来解释不同企业对似乎同质商品定不同的价格，以及对同质的劳动力支付不同的工资的现象。本文正是着重于描述信息不完美条件下的市场均衡问题。我们并不打算给出一个一般性理论，相反，我们将详细描述一个有着重要意义的例子——劳动力市场上的不完美信息。这个例子的某些性质，比如均衡中存在着差异价格（工资）、非自愿失业（劳动力的超额供给）、正利润、多重均衡，以及某些均衡，甚至所有均衡的非优性，具有广泛的适用性，其余的则不然。

本文着重探讨两大类信息不完美现象：

（a）在个人应聘某个职位之前，他并不了解企业将要给付的工资额度，或者他根本就不了解该企业是否有职位的空缺。[1]

（b）对个人而言，除了工资以外，某些重要职位特征，比如工作周的名义长度，只有他在应聘之后才能够知道；其余的职位特征信息（包括同

---

* "Equilibrium Wage Distributions", *The Economic Journal*, Vol. 95, No. 379 (Sep., 1985), 595 - 618. 非常感谢福特基金会、国家科学基金会、洛克菲勒基金会以及古根海姆基金会对本文的资助。S. Salop, J. Mirrless, Barry Nalebuff, P. Diamond, Ian Gale, 以及匿名审稿人的评论让我受益匪浅。本文是 'On the Optimum Wage Structure' (Institute for Development Studies, Nairobi and Cowles Foundation, Yale, 1971) 'Equilibrium Wage Distributions' (Cowles Foundation Discussion Paper No. 375 (1972) 以及 IMSSS Technical Report 154, Stanford (1974) 第一部分的修改稿，自从本文的第一版完成之后，逐渐发展出大量的有关此题目的文献。因此哪怕是一个参考书目我也无法给出。Cowles Foundation 版有更加深入的讨论，其中 Section IV. 3 之后有关于两部门模型的一个扼要的探讨。

1 另外，他们可能知道或者不知道工资的分布情况。在以后的部分我们需要更加明显地假设出关于个人在做搜寻的决策以前所拥有信息的情况。

事的个性）只能够慢慢地了解。

这两大类信息不完美现象存在着重要的区别：在任何经济体系之中，只要能够将个人与其所适合的职位配对，就会有回报。只要存在着个人与职位之间的不匹配，上述信息不完美现象就必然存在；另一方面，存在工资差异的信息不完美则完全是由市场自身造成的。只要自身愿意，一个社会主义经济完全可以选择对同质劳动支付相同的工资。过去，我们常常把市场视为一个很好的信息传递机制，比如传递市场需求和供给。而我们的分析表明：从效果上看，在确定情况下的市场，有可能产生某些不必要的不完美信息，后者很可能会造成非常大的损失。

构造信息不完美条件下的均衡模型存在着两大难题：第一个难题是怎样才能够避免最后信息的累积？

假设个人的生命是无限期的，工资额永远恒定，那么通过搜寻，[2] 最终每个人都能找到最适合他的职位。这样“不完美信息”就只能作为对短期经济特征的一个描述。在下面的讨论中，我们假设市场通过不断地吸纳新的劳动力，以及不断地逐出已经“死亡”的劳动力，保持市场中个人永远存在着“缺乏信息的流动人群”。[3] 这些“缺乏信息的流动人群”数目是如此之大，足以抵消掉通过搜寻所产生的“拥有信息的流动人群”，从而使得不完美信息均衡得以维持。

第二个难题在于如何使不同企业支付不同的工资？因为不完美信息只能够解释为何个人愿意购买高于市场最低价格的商品，或是会接受低于市场最高额度的工资。但它不能解释为何某些追求利润最大化的企业会索取或者支付与别的企业不同的价格或者工资。

我们认为，企业通过支付较高的工资能够减少劳动力的外流，并提高对员工进行专业化培训所带来的期望收益。在确定性条件下，培训成本的上升正好可以抵消低工资带来的好处。要想存在着差异价格均衡，就必然要求作为工资函数的利润具有不止一个相对最大值，而所有在不同的相对最大值点取得的利润最大值都相等。本文让人惊讶的结果是：此结论所要求的条件是非常简单的，甚至在所有企业的培训成本都相同的条件下结论也成立。

---

2 倘若搜寻是有成本的，那么个人将会放弃获得“完美的信息”，从而差异价格（工资）现象得以持续。

3 倘若经济存在着诸如技术变迁等扰动，这样“职位”就具有持续无限期，那么分析方法与这里采用的分析类似。

在追求最高工资的职位与追求最适合的职位之间存在着很重要的联系。差异工资的存在显然会影响到找到更加适合的工作的可能性。而不那么明显的是，如果存在着第二类均衡（“个人与职位相配对”）就不存在单一工资零利润的均衡。

以前的研究侧重于以企业行为的角度来研究个人离职率受工资率的影响（Mortenson，1970；Salop，1973b）以及当存在着差异价格时市场上个人的行为（Salop，1973a）。前者把离职率（quit rate）作为外生变量，而后者把差异工资视为给定的，在这里我们把二者都作为内生变量。

我们的分析可以看作是传统搜寻模型的一个简单的一般均衡基础。这种基础之所以是必要的，有三个原因：

a）如果没有这个基础，我们可能被误导去建立一个根本不存在差异价格的模型；[4,5]

b）如果没有这个基础，我们就无法正确地判断不完美信息条件下的市场行为的福利情况；

c）通过解释为何存在着失业情况下工资不能够下调，它可以成为宏观经济学失业模型的微观基础。

实际上，均衡的具体特征是非常依赖于具体的假设条件的，比如关于生产、搜寻、偏好、劳动力供给以及信息的假设条件。在下一部分我们会给出这个模型的一个最简单的版本。在证明了均衡确实存在着差异工资，以及事实上在差异工资条件下零利润的充分就业均衡之后，我们在第二部分证明，其实还存在着正利润情况下非充分就业的多重均衡。我们还证明了，差异工资的均衡不是帕累托有效的，这些均衡非常不稳定，尤其是可能存在着令人

4 更加一般地说，局部均衡模型很可能是不一致的，比如假定的离职率函数会引致差异工资，而后者又会引致出一个离职率函数，而这个离职率函数可能与假设的并不一致。

5 自从本文最初写成之后，其他的一些价格或者工资差异理论也逐渐建立了。尽管我们无法将这一浩瀚的文献做一个完整的回顾，我们还是可以理出一些重要的脉络：在 Salop（1977）以及 Salop and Stiglitz（1977）的文章中，差异价格是用来对搜寻成本不同各人实行歧视政策。在本文早期的版本以及 Reinganum（1979）的文章中，证明了采用不同技术的企业会支付不同的工资或者索取不同的价格。在 Mortenson（1973）的文章中证明了，搜寻成本的存在使得劳动力市场无法对于次级市场的扰动采取完全套利的反应（这个分析类似于 Grossman and Stiglitz（1976）的分析，他们证明了信息成本的存在阻碍了资本市场达到完全的套利）。在本质上最接近本文的是那些试图证明即使在同质的个人和企业以及没有外生扰动源的条件下，均衡可能是存在着差异价格的一类文献（Salop and Stiglitz，1982；Butters，1977）。在巴特斯（Butters）的模型中，索取低价的企业单位广告成本下可以吸引更多的客户，同样在我们的模型中，支付高工资的企业能够留住员工更长的时间。

预想不到的比较静态性：伴随着搜寻强度的增加，工资的差异程度也会增加。第三部分我们证明了，这些核心的结论在扩展的条件下（即搜寻强度和培训成本是内生条件下）依然可靠。最后在第四部分我们证明了，如果企业具有某些只有在培训结束后才能发现的非金钱（non-pecuniary）特征，那么唯一的均衡就可能保证就业的存在。

一、基本模型

1. 引言

本部分的模型是传统搜寻模型的一个简化版。模型的基本要素（下文有更加详细的描述，并对要素进行了不少修改）如下：

(a) 假设个人不断地搜寻更好的工作，在简化模型里，这意味着更高薪水的职位。一旦他找到了，那么他会立刻离开原来的职位（在基本模型里，搜寻的强度是外生给定的参数）。

(b) 为简单起见，假设只有一种商品，我们把它视作基本计价单位。假设所有的生产都服从恒定的规模收益，并且市场是自由进入的。各个企业有着不同的培训成本。通过支付高工资，企业能够降低员工的离职率。最优的工资水平依赖于企业的培训成本。

关于自由进入市场，以及技术的规模收益不变的假设使得零利润的条件看起来很自然。这又意味着企业的规模不是决定因素，真正的决定因素是工资差异。之后我们将会用几个例子证明零利润条件可能是不满足的。

我们使用的模型实际上是动态模型。企业在做出决策的时候必须对未来的利率、价格及其他的市场条件做出预期。我们回避了对于长期均衡的关注，那里的工资分布、价格等都是不变的。[6]

均衡条件下的每个企业都具有某一特定的工资、就业和产出水平，工资分布是市场均衡条件下的特征。在均衡条件下：[7]

(a) 给定离职率（由工资分布决定），每个企业会选择最优的工资水

6 由于我们的模型具有随机性，所以为了避免总体变量产生随机波动，我们假设经济体是很大的。

7 在更加一般的模型之中，为了定义均衡还必须确定企业的技术选择、个人的搜寻强度选择，以及不同商品的相对价格。相应地，模型中还有其他均衡条件，比如每一个企业都选择某技术，从而实现利润以最大化；每个人都通过选择某一搜寻强度，实现其期望效用的最大化（给定他现在所在企业的工资水平、搜寻成本以及工资分布）。

平，以便实现每个工作（职位）利润的最大化。

(b) 所有职位的利润为零。

(c) 劳动力和商品市场是出清的。

在部分的其余内容，我们将通过模型分析简单经济体的均衡情况。分析过程分为以下几步：在第2点我们描述了经济体中均衡时的个人行为；在第3点我们描述了在给定离职率的情况下的企业行为；在第4点我们推导出了离职率函数；第5~8点分别构造出了单一工资、双重工资、多重工资以及连续工资条件下的均衡。

2. 个体行为

假设所有人是同质的。[8]他们的死亡率μ服从指数分布，他们会被等量的工人补充，因此劳动力供给水平保持不变为I。每个劳动力都只能提供一个单位的劳动。当工人进入了劳动力市场，他们应聘并接受某个职位是随机事件。同时他们还会继续搜寻更高薪的职位，从一个企业转到另一个企业，工人发现工资水平所需要的时间是一个服从泊松分布的随机变量。平均每单位时间的搜寻数量s是一个常量。对于任何小于s的搜寻强度，搜寻和转换职位是没有成本的；而大于s的搜寻强度是不允许的。工人并不知道任何一个企业的工资水平。由于s实际上是外生决定的，我们不需要再假设个人是否知道工资的分布。不过，以后（第四部分）我们将会更加明确这个假设。

3. 企业行为：给定离职率函数条件下工资的决定

为了简单起见，假设生产过程仅仅需要劳动力，不需要资本，并且生产的规模报酬不变。因此生产函数可以通过每个工人的固定培训成本，以及每人每年的产出流量来描述，假设前者从一开始雇佣工人之时起就会产生（后面所关注的模型中企业都无法决定技术的选择，他们唯一能够做出的选择就是工资水平。如果企业得到零利润，那么他们对于企业的生产规模选择就是无差异的；一旦企业决定了工资政策，他就会接纳所有应聘该职位的求职者）。企业选择使得每个工人利润现值最大化的工资水平。这些约束条件使得随机均衡分析得到大大地简化，因为此时工资是个常数，并且满足最大化条件

$$P_i \equiv a(T_i) - \{w_i + [q(w_i) + r]T_i\} \tag{1}$$

---

8 在第四部分，我们假设由于对于某个企业的非特别特征具有不同的评价，个人之间存在着差别。同质人的假设是为了避免出现均衡条件下的工资分布，是由于企业为了针对不同个人实施垄断性的价格歧视的情况。参见 Salop and Stiglitz（1977）。

其中 r 是利率，P_i/r 是利润额的现值；$q(w)$ 是离职率函数。劳动力成本包含直接的工资支付（w），培训工人以代替离职的工人的成本（q），以及以前为了培训工人所花费的利息（r）。假设（事实也正是如此）企业必须承担一部分培训成本。阿诺特（Arnott）和斯蒂格利茨（1982）证明了只要工人比企业更加厌恶风险，[9] 那么在最优合同之中这就是合理的。离职率 q 实际上起到了企业对于"人力资本"投资的贴现因子的作用。通过提高工资水平，尽管直接的劳动力成本增加了，但是跳槽成本（"贴现率"）却降低了。当下式满足的时候利润达到最大化（平均每个工人的劳动力成本最小化）：

$$1 + q'(w)T = 0 \tag{2}$$

离职率函数一般情况下是凸函数，如图 1（a）。等成本线是一条斜率为 $-1/T$ 的直线，在与离职率函数的交点处达到最优工资率。当培训成本增加，w 也会平滑地增加。

当然，目前我们并没有给出解释为何离职率函数一定要是如图 1（b）而不是其他的形状。在图 1（b）之中对于 $T<\hat{T}$ 有 $\omega \leqslant \hat{\omega}$。当 T 增加到大于 $\hat{T}$ 的时候，工资会有一个向上的跳跃，此后随着增加而平滑增长。

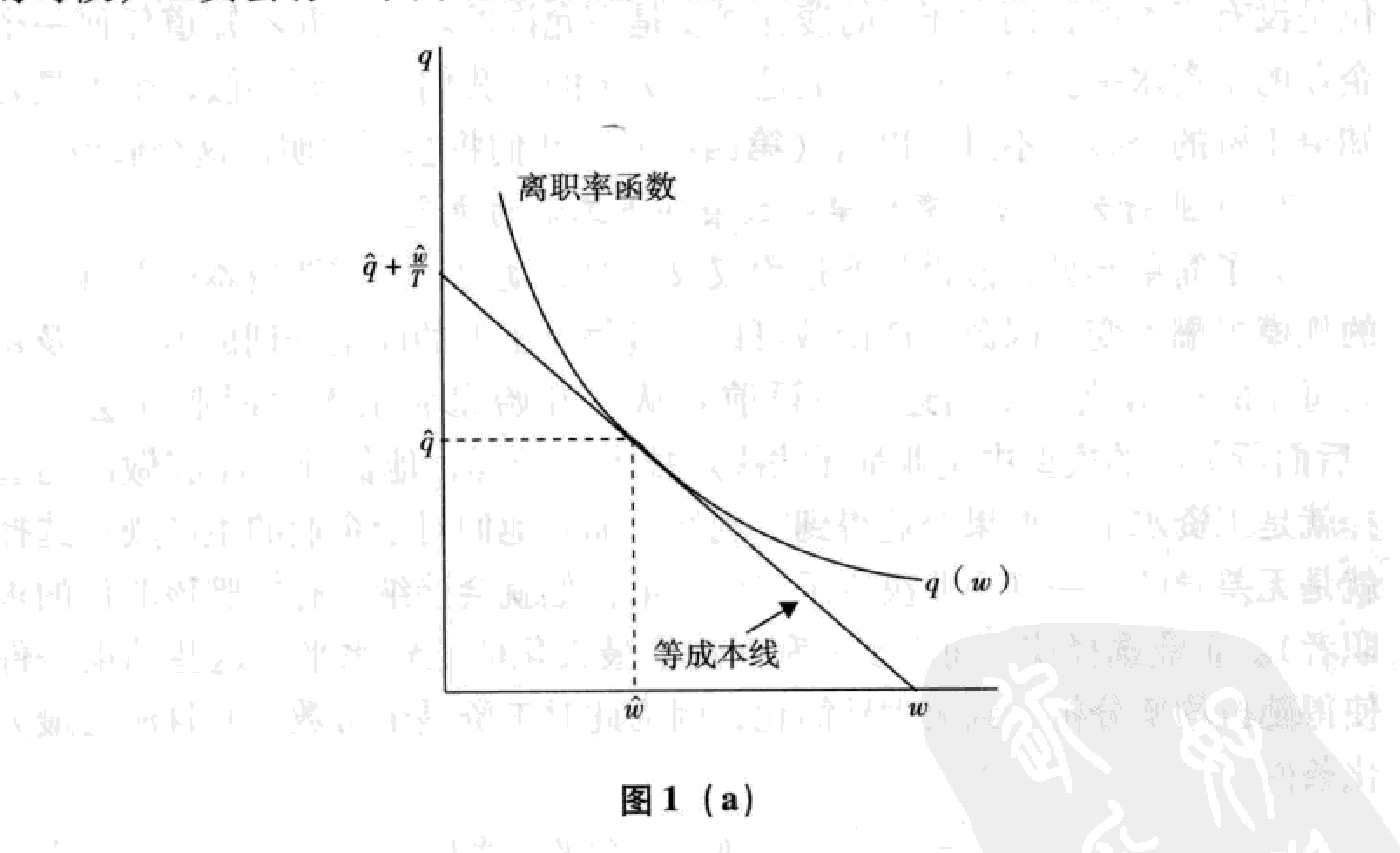


图 1（a）

9 他们的模型假设企业存在着非独有特征（将在第四部分讨论）；不过其结论是非常有一般性的：只要存在着与追求更高工资水平无关的跳槽行为，当然这是很难确认的，那么厌恶风险的工人会希望企业能够承担一部分的跳槽成本。

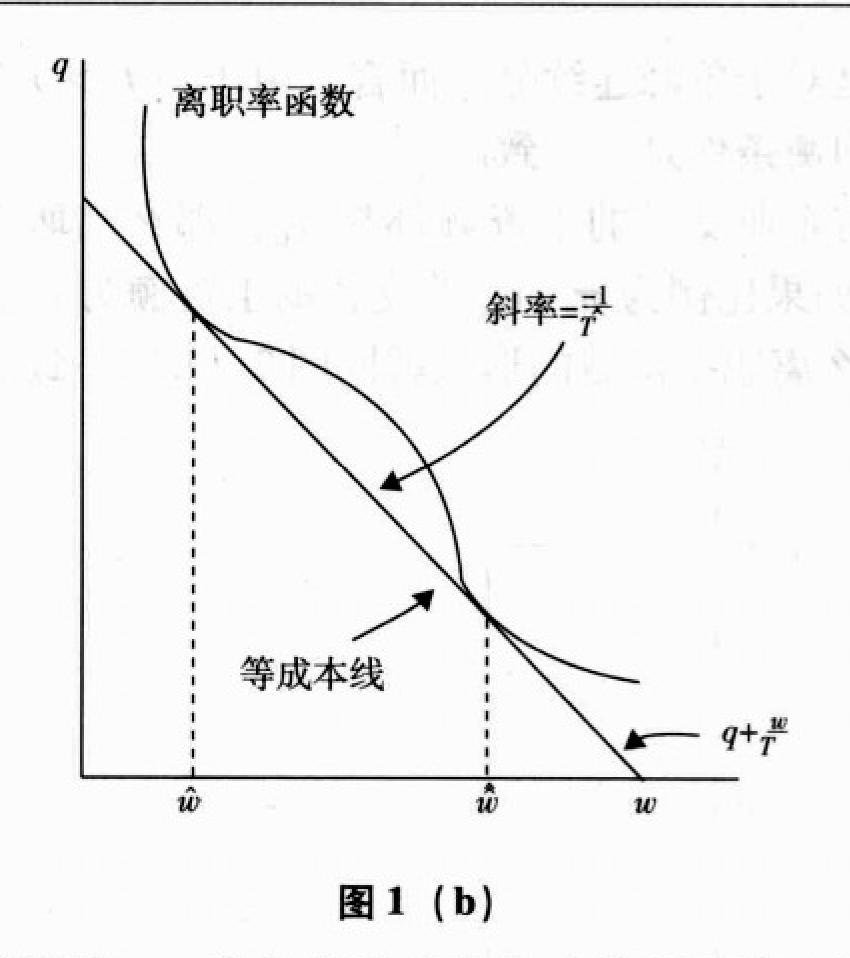


图1（b）

但是在两种情况下，工资都是培训成本的单调递增函数。

4. 离职率函数的决定

现在关键的问题在于离职率函数是如何决定的。在假设条件下离职率应该等于死亡率加上工人找到更好的工作的概率。后者等于单位时间的平均搜寻数，乘以某次搜寻职位能够提供高于现在供职的企业所提供工资额的可能性。为简单起见，假设找到高于现在供职的企业所提供的工资额 w 的职位的可能性，等于市场上支付工资高于 w 的企业所占的比例。[10]如果 F（w）是支付工资水平小于或者等于 w 的企业比例，那么离职率为：[11,12]

$$q(w) = \mu + s[1 - F(w)] \tag{3}$$

于是如果是 w 的可微函数，则

$$q'(w) = - sf(w) \tag{4}$$

其中 f（w）是 w 的分布密度函数。为了满足离职率函数为凸的条件，密度函数必须是单调下降的；也就是说，唯一可能的连续工资密度函数应该是单

10　这个假设与其他假设是一致的，即每一个大企业都有一些职位待聘，而职位的数目与现有工作职位数成比例。

11　如果转而假设另外搜寻一个企业所需的时间是个随机变量，由于我们已经假设了工人每单位时间搜寻的企业数为 s，那么 q（w）$=\mu+\{1-[F(w)]^s\}$，$q'=-sF^{s-1}f$。其余部分的计算方法完全可以通过类似的修改直接得到。

12　由式（3）显然可以看出，离职率函数不仅仅是该企业支付工资水平的函数，还是其他所有企业支付工资水平的函数。因此，如果存在着 m 个企业，我们可以把第 i 个企业的离职率写为 $q_i=q_i$（w_1，w_2，$\dots w_i$，$\dots w_m$）。此种标号方式要求与 q_i 是依赖于 w_j 的，而 $\partial q_i/\partial w_i$ 计为 q'（w）。

调递减的。特别是对于单峰连续分布而言，由于当$f' > 0$时工资分布不满足该条件，所以与均衡条件是不一致的。[13]

如果其他所有企业支付的工资额都是w_0，那么离职率函数的形状就同图2（a）一致；如果比例为π的企业支付的工资额为w_1，其余企业支付的工资额是w_n，那么离职率函数的形状就同图2（b）一致。培训成本为

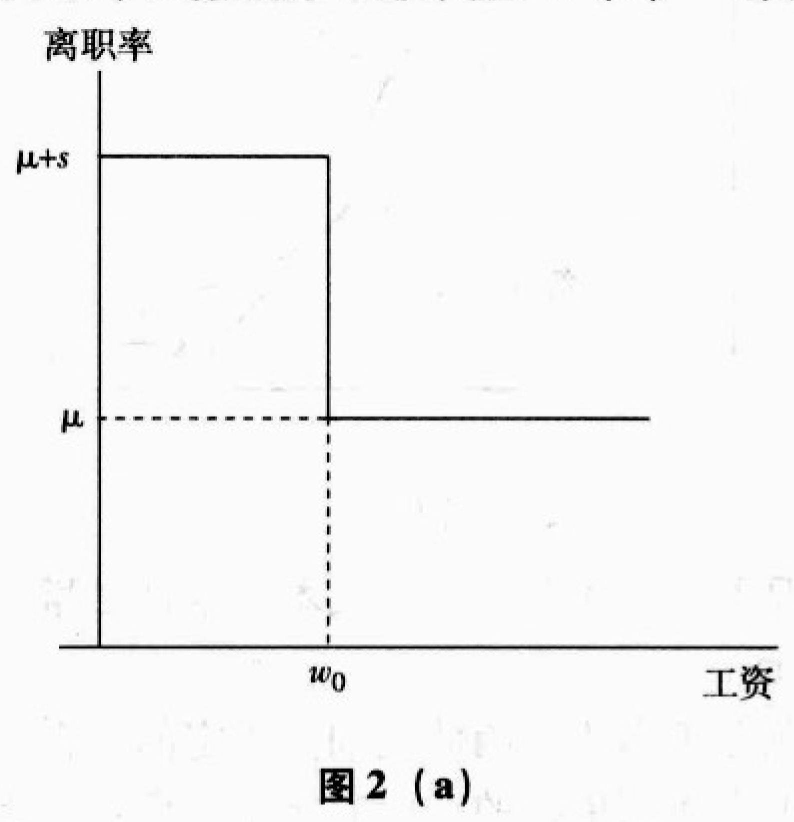


图2（a）

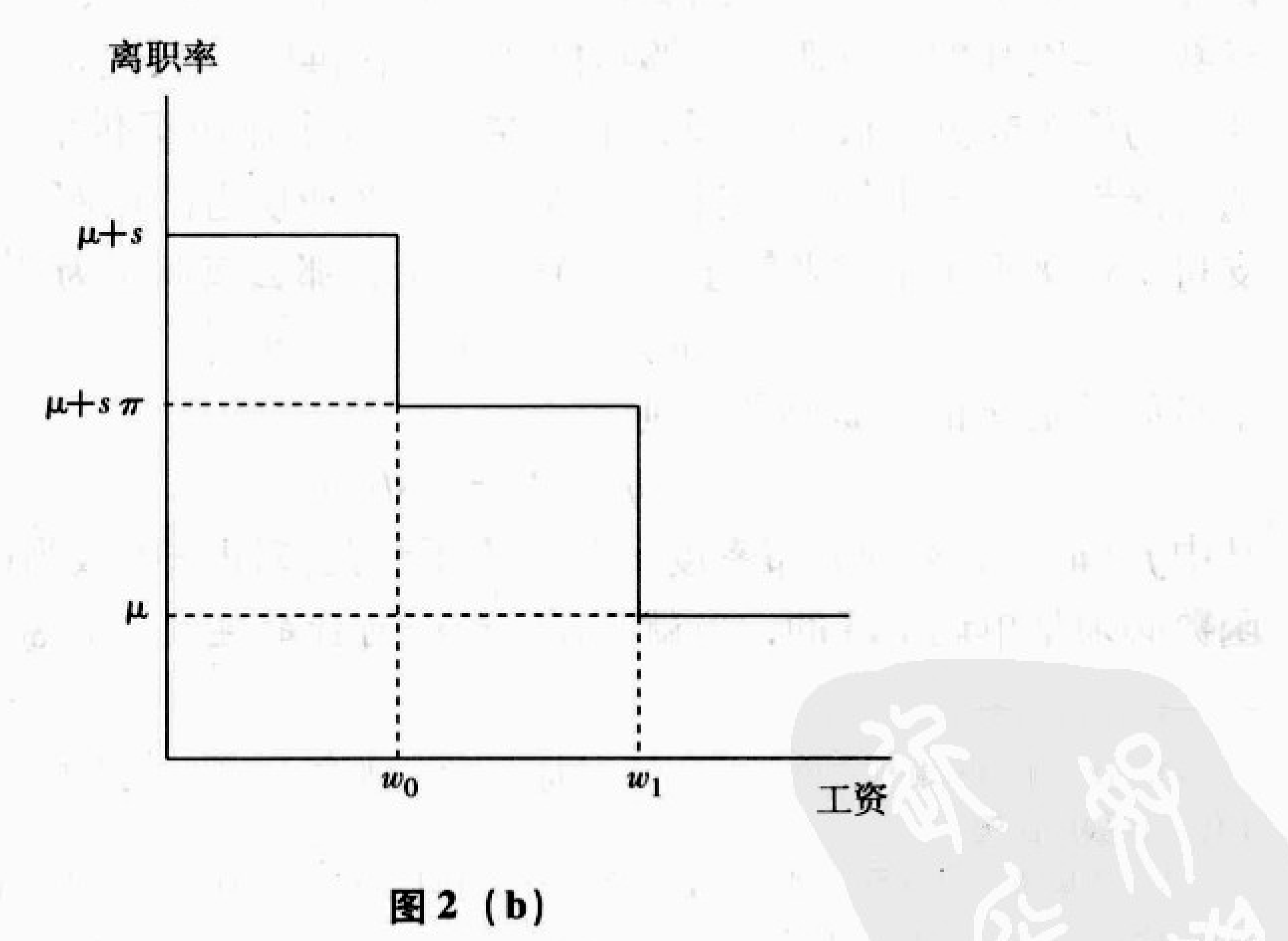


图2（b）

13 对于此结果不必过分强调，因为，正如我们所见到的情况，当模型更加复杂后结果就不成立了。

$$T = \frac{w_1 - w_0}{s\pi} \equiv \hat{T} \tag{5a}$$

的企业对于支付 w_0 或者 w_1 是无差异的；培训成本高的企业支付 w_1，培训成本低的企业支付 w_0。同理，如果比例为 π_1 的企业支付的工资额为 w_1，比例为 π_2 的企业支付的工资额为 w_2，离职率函数的形状会类似于图2（c），而培训成本为

$$T = \frac{w_2 - w_1}{s\pi_2} = \hat{T}_2 \tag{5b}$$

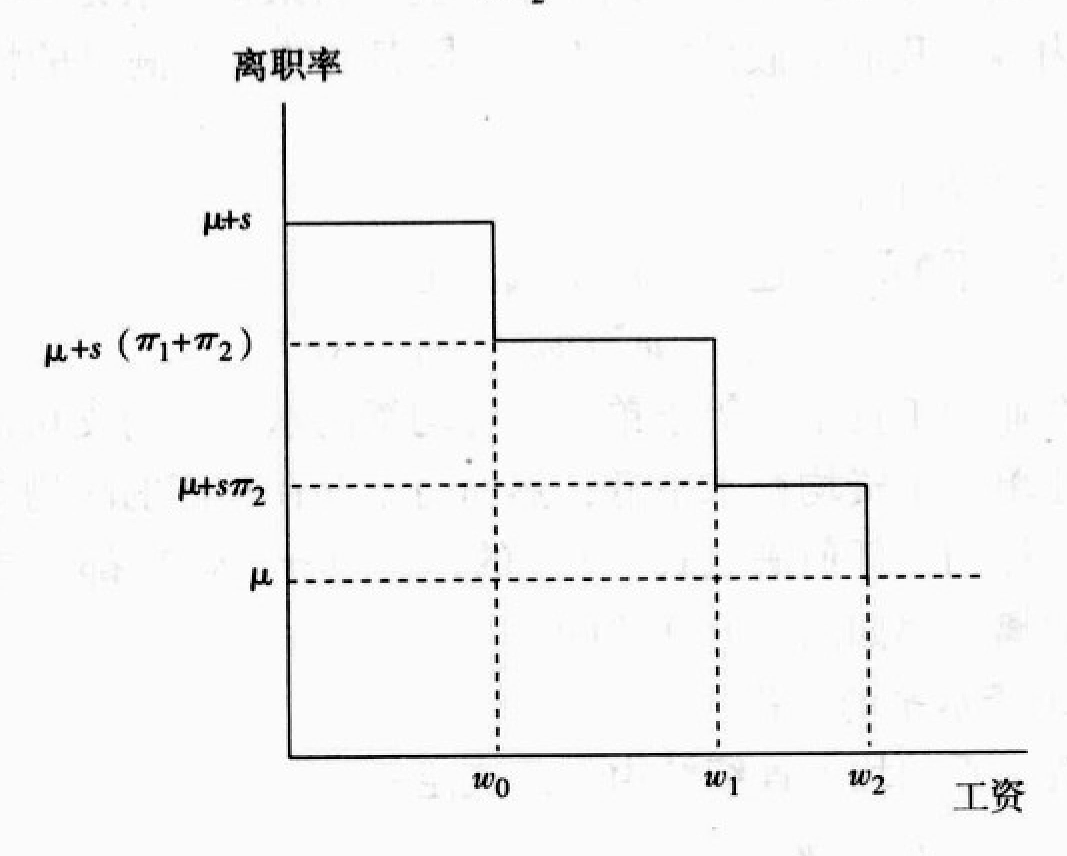


图2（c）

的企业对于支付 w_2 或者 w_1 是无差异的；而培训成本为

$$T = \frac{w_1 - w_0}{s\pi_1} = \hat{T}_1 \tag{5c}$$

的企业对于支付 w_0 或者 w_1 是无差异的；培训成本高于 max［$\hat{T}_2$，（$w_2 - w_0$）/s（$\pi_1 + \pi_2$）］的企业支付 w_2，低于 $\min[\hat{T}_2, (w_2 - w_0)/s(\pi_1 + \pi_2)]$ 的企业支付 w_0。

注意到如果有企业选择了 w_1，$\hat{T}_1 < \hat{T}_2$，即

$$\frac{\pi_1}{w_1 - w_0} > \frac{\pi_2}{w_2 - w_1} \tag{5d}$$

在连续工资分布密度函数应该是 w 的单调减函数这个条件下。我们得到的结果是离散的。

5. 单一工资水平的均衡

单一工资水平的均衡可以由零利润条件得到

$$a(T) = w^* + (\mu + r)T \tag{6}$$

需要假设如果 w^- 是工人可以得到的最低工资[14]

$$w^- + (\mu + r + s)T > w^* + (\mu + r)T$$

$$s > \frac{w^* - w^-}{T} \tag{7}$$

由于搜寻强度是如此之高，以至于企业支付最低工资是不划算的。在本部分的余下内容，我们都假设式（7）式是满足的，不满足的情况留待下一部分讨论。

6. 双重工资水平的均衡

双重工资水平的均衡连续统由下式决定

$$w_1 = w^*, w_0 = w_1 - s\pi T \tag{8}$$

支付高工资企业的工资水平等于单一工资均衡的水平，而支付低工资企业的工资水平，比单一工资均衡水平低；然而由于存在较高的跳槽率，他们的劳动力成本是一样的。任何满足式（8）的组合（π，w_0）都构成均衡。而此均衡条件下的离职率如图 3（a）所示。

7. 三重工资水平的均衡

三重工资水平的均衡连续统由下式决定：

$$\left.\begin{aligned} w_2 &= w^* \\ w_1 &= w_2 - s\pi_2 T = w^* - s\pi_2 T \\ w_0 &= w_1 - s\pi_1 T = w^* - s(\pi_1 + \pi_2)T \end{aligned}\right\} \tag{9}$$

任何满足 $\pi_1 + \pi_2 < 1$ 以及式（9）和不等式（5d）的组合 $\{w_1, w_0, \pi_1, \pi_2\}$ 都构成均衡。

8. 多重工资水平的均衡

显然，我们能够构造出任何有限数量工资水平的均衡。实际上，可能在 π^* 和 w_0 之间存在着均匀密度分布为 $1/sT$，聚点位于 w_0，并且满足 $w_0 > w^* - sT$的均衡工资水平连续统（如图 3b 所示）。

14 对于 w^- 还可以有其他的理解。在存在着有效的最低工资法案的经济之中，w^- 就是那个最低工资。另外 w^- 还可以视为保留工资水平，低于此水平的话工人就不会工作了。或者还可以把 w^- 理解为工人们自我雇佣的工资水平。另外，经济之中可能还存在着某个组织化的完全竞争劳动力市场，由于效率工资产生了最低工资（见脚注 17）。

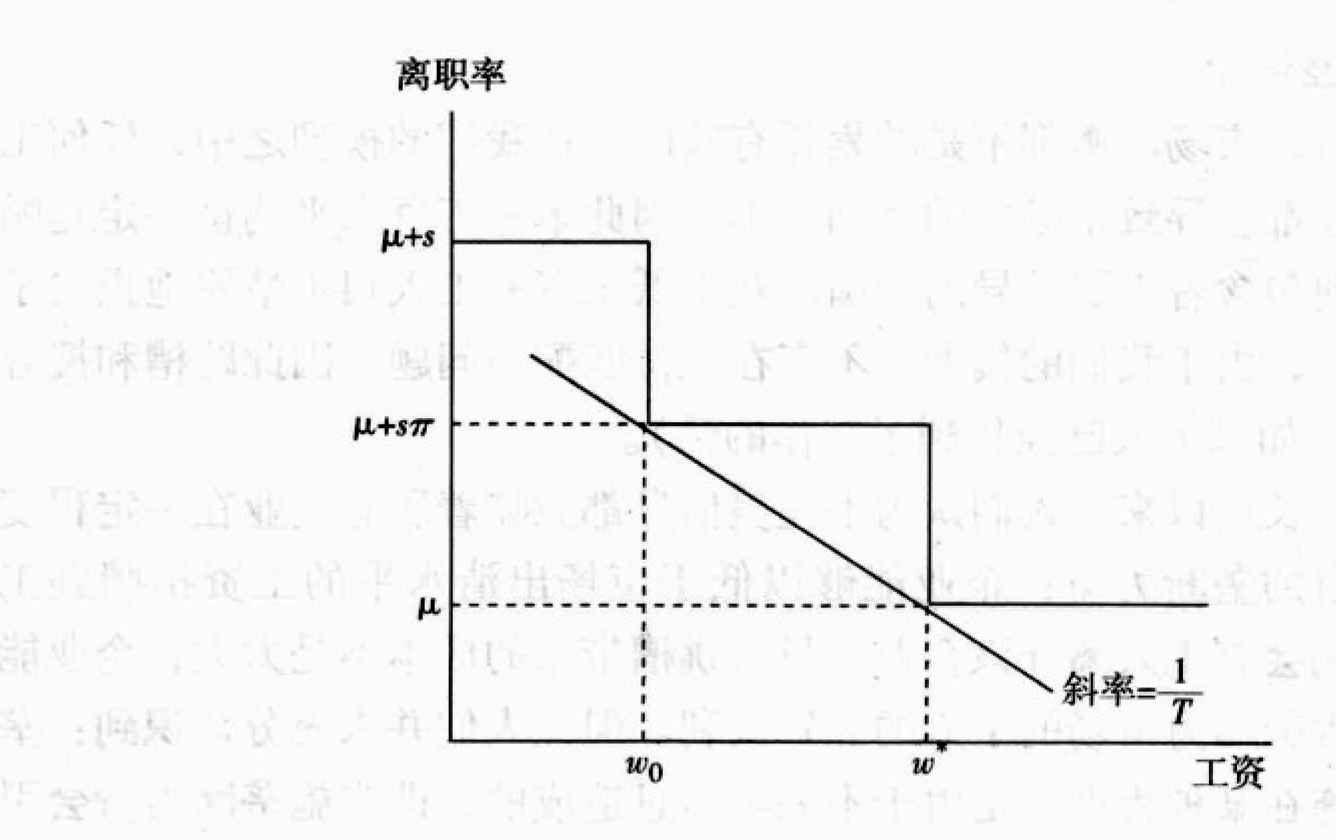


图 3（a） 双重工资水平均衡

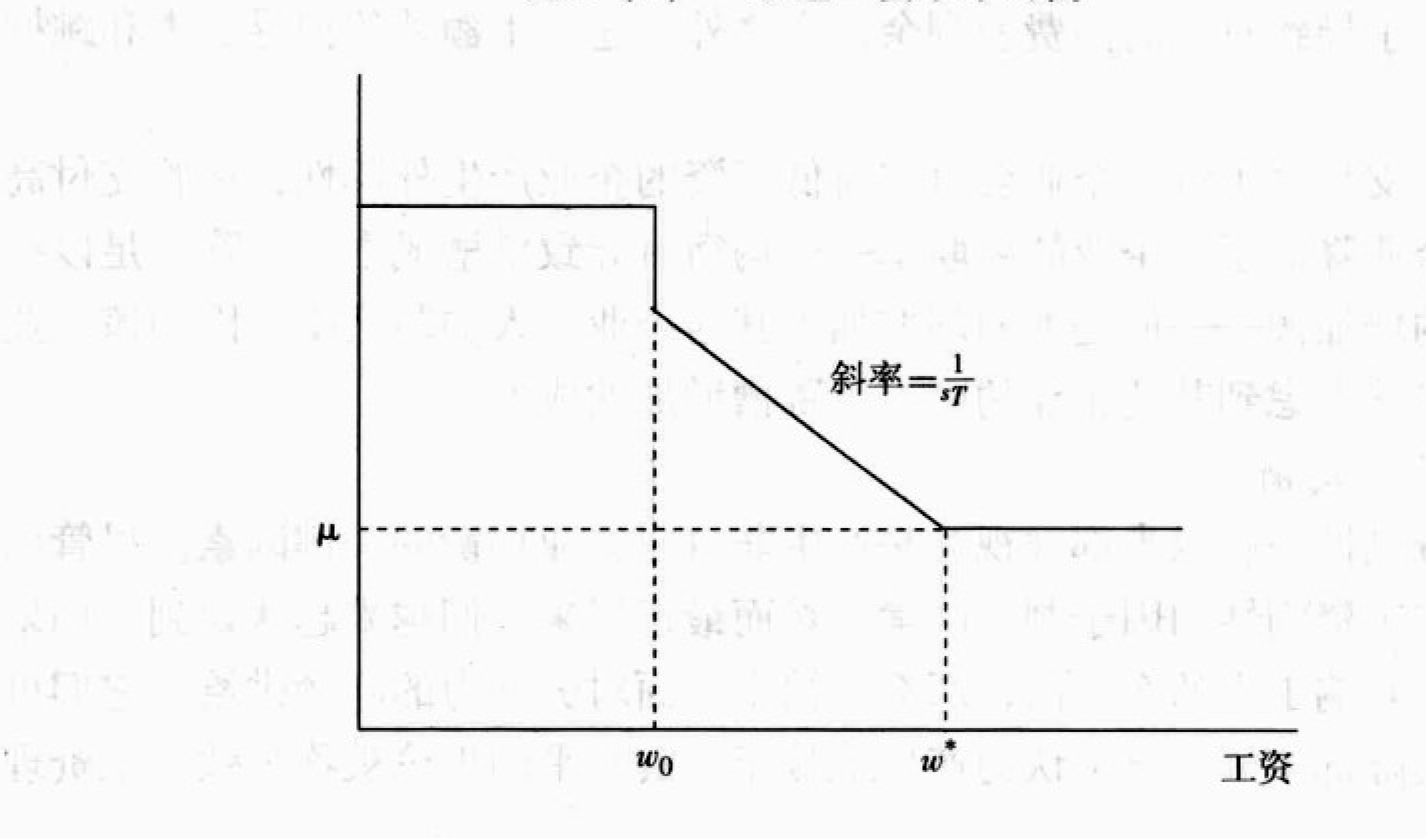


图 3（b） 聚点为 w_0 的连续工资分布的离职率函数

二、基本模型的含义

下面我们要探讨这个模型有的一些重要含义。首先，我们要证明大部分均衡都不是帕累托有效的。其次，我们将要证明，对模型稍稍进行一点改动，就能够产生出非自愿失业以及正利润。最后，我们要考察结果对于搜寻均衡的比较静态分析的意义。

1. 福利经济学

一般而言，市场均衡都不是帕累托有效的：在我们的模型之中，任何工资水平的变迁都会导致不必要的培训成本。因此单一工资水平均衡一定是帕累托优于任何包含着工资差异的均衡。获得低工资的工人很不情愿地损失了福利。实际上，由于我们的模型里不存在职位匹配的问题，因此跳槽和搜寻是不必要的（如果工人已经找到了工作的话）。[15]

可以说，长期以来，人们认为不完美信息都意味着所有企业在一定程度上都具有各自的垄断力量；企业能够以低于市场出清水平的工资招聘到工人。虽然立刻会有工人离开该企业，只要跳槽带来的成本不是太大，企业能够利用新进入劳动力市场的工人的无知获利。但是人们并未充分认识到：垄断力量的实施在某种程度上是由于不完美信息造成的，即实施垄断力量会引致工资差异，而如果没有工资差异的话，搜寻就是不必要的了。因此运用垄断力量除了导致直接的消费者剩余损失之外，还产生额外的搜寻成本和跳槽成本。[16]

由于支付高工资的企业会对支付低工资的企业产生外部性：尽管支付高工资的企业降低了本企业的离职率——均衡时导致跳槽成本的下降，足以抵消工资的增加额——但是他们却增加了其他企业工人的跳槽率。因为该企业决策时未能考虑到其他企业的因工人跳槽增加的成本。

2. 失业均衡

工资刚性一直以来都被视为是产生非自愿失业均衡的关键因素。尽管旧的文献将工资刚性归因于制度因素，然而最近以来人们越来越认识到，如果工资能够影响工人的净产出，那么哪怕企业面对劳动力的超额供给，它们可能也不会降低工资。此类认为产出依赖于工资水平的理论被称为效率工资理论。

通常人们在回答为何不存在非自愿失业的时候会说，如果有工人失业了，那么他们可以提出降低工资条件来换取工作；如果工资水平被不断地降低，劳动力的需求水平就会上升而供给水平就会下降。整个过程只有当供求相等的时候才结束。在效率工资模型之中，这种说法是不成立的：如果降低

---

15 在目前所给出的模型之中，实际上根本就不存在搜寻的成本；可以很直接地将模型扩展到存在搜寻的成本的情况下。见下面的第三部分第 2 点。

16 实际上，如果补偿的劳动力供给弹性很小，那么搜寻成本和跳槽成本就是实施垄断力量所造成的福利损失的主要来源。

工资会导致利润下降的话，那么企业就不愿意降低工资水平了。提高工资水平可能会提高产量水平，因为高的工资可以使得工人更加努力工作（Calvo，1979；Shapiro and Stiglitz，1984），而且企业可以提高求职者的总体素质（Stiglitz，1976，1982；Weiss and Stiglitz，1980；Nalebuff and Stiglitz，1984），或者增加求职者的总体人数，这样企业就能够选择那些更加适合该职位的工人。[17]

降低工资水平增加了跳槽率，因此即使存在着要价更低的工人，企业不会降低工资水平。

关于这点有一个很自然的问题：如果工人无法通过降低工资水平来获取就业，那么为什么他不提出承担大部分（甚至全部）跳槽成本来获得就业呢？关于这一点至少有三种可能的解释：第一，工人可能不具备支付培训成本的能力。第二，企业有道德风险的问题：企业可能会收取申请费，说是为了培训工人用，等到收到钱以后就炒掉工人。第三，愿意花最大代价获得某个职位的人，也许并不是最适合该职位的人（这里存在着质量选择效应）。对上述问题的证明方法类似于魏斯和斯蒂格利茨关于对合同双方进行约束可能无法实现信贷市场的均衡，以及夏皮罗和斯蒂格利茨关于仅凭担保无法消除激励与失业的两难境地的证明。

关于不应该向工人收取部分培训费用的论证，可以用来反驳所有收取进入费用的行为。向自愿离开的个体收取离职费用也面临类似难题：一方面，它会促使企业提供较差的工作环境，以至于工人会选择跳槽（于是企业得到了离职费）；另一方面，找到更好职位的工人会倾向于偷懒，因为这样会促使企业将其炒掉（从而避免向企业支付离职费）。除非个人提供担保，否则，企业是很难收取到离职费的，但是由于工人的资本量有限，因此担保是不太可信的。[18]如果工人的资本量有限，那么均衡中合同必定包含工资结构以及差异补偿[19]（同离职费的功能相似），但是在合同最初的时期内，企业

---

17 尽管这些理论的历史非常久远（至少可以追溯到马歇尔），但是对它们赋予现代意义的第一人，应该是 Leibenstein（1957），他将产量依赖于工资，归因于对工人的培养。Akerlof（1984）从社会心理学的角度解释了产量为何依赖于工资。近来的发展见 Stiglitz（1984）和 Yellen（1984）。

18 在别的文献里，对于合同条款设置限制被证明会对企业行为产生其他的种种影响，见 Grout（1984）与 Crawford（1983）的著作。

19 在以前的一篇文章“On the Optimum wage structures”（1971）（那是本文经过扩充的修改版本）中，我分析了最优的工资结构如何随着时间而变化。我们的假设同 Shapiro and Stiglitz（1984）中的假设类似，暗含假定静态最优的工资结构的存在，但是在更加一般的模型里，这个假设可能是不准确的。

仍然要承担一部分的培训成本。

因此，随时间递增的工资结构是无法完全解决由强迫工人承担培训成本所造成的企业道德风险以及选择效应难题的（由于跳槽成本是浪费性的，所以一旦有企业向工人提供更加高薪的职位，工人与所供职的企业之间就得讨价还价一番。而谈判的结果有赖于一系列因素，包括双方的贴现率，以及他们所拥有的信息。在保持模型假设一致性的同时，我们可以稍作修改，认为企业无法确认到底工人是否已经有了新的东家，以及如果真的有了，工资额为多少）。

因此，尽管均衡条件下的合同会在某种程度上比我们假设的情况要更加复杂，但是对于分析结论而言，真正的重要性在于：出于某些原因，企业会承担部分人员更替成本，而人员更替成本会受到合同条款的影响，正是由于这个原因，就算是在失业条件下，修改合同条款，在降低工人福利的同时，很可能也会降低企业的利润。

为了显示均衡失业的性质，我们将规模报酬不变的假设改为规模报酬递减；我们把总体生产函数定义为 F（L），充分就业条件下的工资水平满足下式

$$F'(\bar{L}) = w^f + (\mu + r)T \tag{10}$$

$\bar{L}$ 是全部可用的劳动力数量。充分就业水平条件下边际劳动产品的值等于充分就业条件下的工资水平加上跳槽成本。再假设所有企业支付的工资水平 w^f 都完全相等，并且

$$w^* > w^f \tag{11}$$

这样，企业仅仅希望在如下条件下雇佣工人

$$F'(L^*) = w^* + (\mu + r)T \tag{12}$$

$$L^* < \bar{L} \tag{13}$$

也就是，存在着非自愿失业的条件。只要满足条件（7），就不存在能够令工人满意同时利润超过 w^* 的工资水平。因此企业既没有动力改变工资水平，也没有动力改变雇佣工人的数目。

注意到如果所有企业能够同时改变支付的工资水平，那么经济就能够实现充分就业。因此非自愿失业源于过高的工资，但是过高工资的产生，并不是由于工会的力量，而是不同企业之间的劳资政策互相影响以及他们之间无法相互协调的结果，这就造成了无效率的纳什均衡。

实际上，在这个模型里很容易看出来扰动是如何把经济从充分就业的均衡移动到非充分就业的均衡。假设最初经济位于充分就业的均衡处，但是技

术存在着扰动从而降低了边际劳动产品的值。只要其他所有企业支付的工资水平为 w^*，那么企业采取如下策略就是有利可图的：所有的调整行为都只对已就业工人有效。[20]如果所有企业都相信其他企业会在价格变化条件下，维持名义工资水平不变，那么该企业维持名义工资水平不变就是有利可图的。该模型与那些名义刚性和实际刚性模型是一致的。

尽管在我们建立的模型里存在着充分失业的单一工资的均衡，在第四部分，我们将证明单一工资的均衡有可能引致非自愿失业。

3. 正利润均衡

尽管看起来对于自由进入条件下，对规模收益不变模型设定零利润条件似乎是合情合理的，但是在存在搜寻成本的模型中可能存在着正利润均衡。通常我们认为，之所以均衡时的利润为零，是因为假如利润为正，那么一些企业就会从别的企业那里将工人挖过来，从而提高工资水平；这个过程会一直持续下去，直到企业利润为零为止。但是提高一点工资水平并不足以保证马上就能够招聘到工人，除非搜寻成本为零。这样，通过稍稍地提高工资水平，企业能够招聘到多一点的工人——这些工人的工资水平高于该企业原来的工资水平，但是低于现在给出的工资水平。[21]该企业必须在提高招聘到工人的可能性与降低每个工人的利润之间做出选择。均衡时要求工资水平必须使得企业利润达到最大化，但是并不一定为零。在均衡中，[22]所有的工资水平一定产生相同的期望利润水平，而且该期望利润水平值要大于其他的工资水平下所达到的期望利润水平。由于申请某职位的工人接受该职位的可能性为 F，工资分布函数必须满足

$$\left.\begin{aligned} RF &= k,\text{对于所有已支付的工资水平} \\ RF &\leq k,\text{对于所有未支付的工资水平} \end{aligned}\right\} \tag{14}$$

其中 R 是从每个工人身上得到的期望利润的现值，它满足

$$R \equiv \frac{a - w - [\mu + s(1 - F) + r]T}{r + \mu + s(1 - F)} \tag{15}$$

---

20 技术的其他变化，比如 s，T，r 或者 μ 的变化同理。注意到真实利率的增加在本模型之中会导致非自愿失业。

21 这里我们排除了企业通过其他的方式招聘工人可能性，比如通过广告等。而这些方式也不是没有成本的。我们的假设保证了某企业招聘到的工人数目不依赖于它所支付的工资水平。不过当信息结构的假设改变的时候，这个结果不再成立。

22 假设企业数目是连续统。

这里给出的工资分布函数形式如图 4 所示。[23]

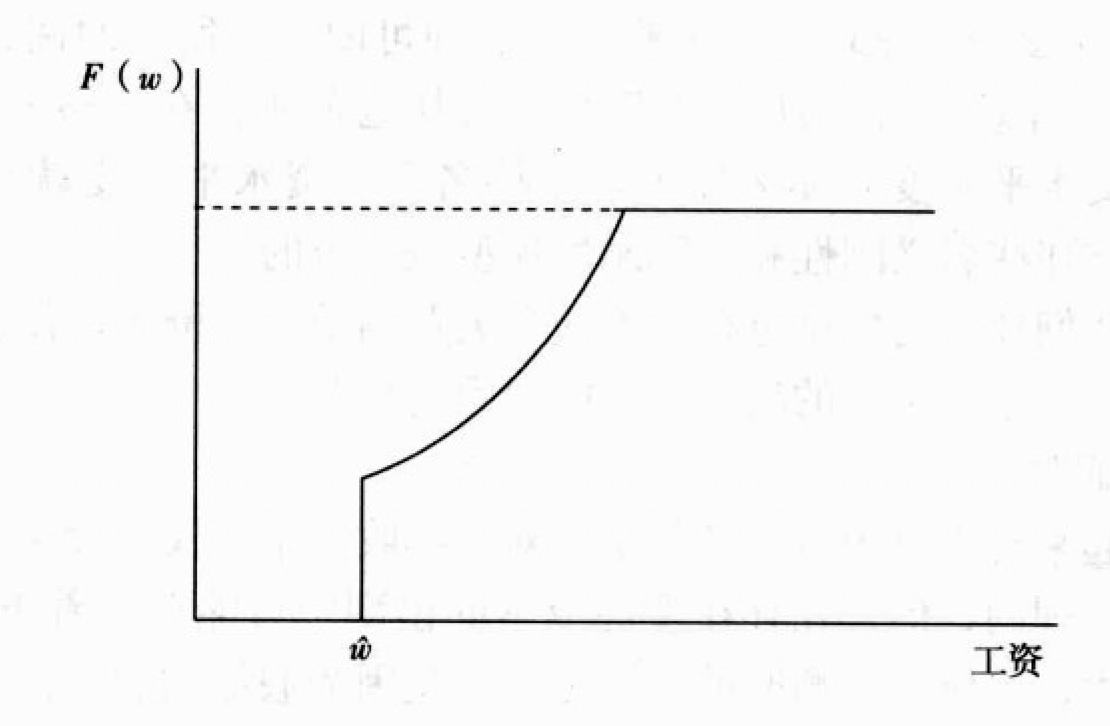


图 4

尽管到目前为止，我们还没有给出企业数目（职位数目）的决定方式，不过只要对模型略作修改就能够决定它。假设每个职位需要有一台价值为 K 美元的机器，在式（15）所定义的就是期望准租金的现值。当有企业进入的时候，申请到某个企业的工人数目必定下降，并且会导致工人离职率的上升，职位空缺能够填满的期望时间增加了；而只有当期望准租金的现值等于

---

23 确切地说，工资分布函数形式服从

$$F = \begin{cases} \dfrac{w - a + (\mu + r + s)T - sk + \sqrt{\{[w - a + (\mu + r + s)T - sk]^2 + 4sTk(r + \mu + s)\}}}{2sT} & \text{对于 } w \geqslant \hat{w} \\ 0 & \text{对于 } w < \hat{w} \end{cases} \tag{16}$$

注意到所有均衡都必须满足式（16）。由于还不能决定 $\hat{w}$ 和 k，所以还是存在着多重均衡。在正利润条件下是不可能在最高工资处达到聚点均衡。因为如果那样的话，稍稍增加工资水平就能够导致工人接受职位可能性发生不连续变化，同样利润水平也会发生不连续变化。在图 4 之中，在 $\hat{w}$ 处存在一个聚点均衡。尽管在高工资处利润水平大于零，稍稍增加工资水平并不能招聘到更多的工人，也不会降低离职率；因此企业没有提高工资水平的激励。

机器的价值的时候，才会有企业进入市场。[24,25,26]

4. 比较静态分析

由于均衡的连续性，很难进行比较静态分析。我们特别关注的重点在于搜寻强度的增加是否会对工资分布函数产生限制？尽管局部均衡分析的结果说明，会产生限制，但是本文所作的一般均衡分析结果说明完全相反的结果也是可能的。考虑一个双重工资水平的均衡，s 的增加要么降低支付低工资的企业的工资水平——因为当跳槽率增加以后企业不能够相应地提高工资水平——要么会降低支付高工资的企业的比例（只要 $s\pi$ 是常数，支付低工资的企业的跳槽率和支付高工资的企业相同）。

5. 均衡的稳定性

本文着重考察了均衡分析；然而多重工资条件下的均衡明显地具有不稳定性是值得关注并且富有启发意义的。考虑一个二重工资条件下均衡，如果某个企业突然将其工资水平从 w_1 变为 w_0，而 w_0 比 w_1 能够带来更高的利润；那么所有其他的企业都会改变工资水平。同样，如果某个企业突然将其工资水平从 w_0 变为 w_1，而 w_1 比 w_0 能够带来更高的利润；那么所有其他的企业也都会改变工资水平（因为此时的人员更替率降低了）。这个单一价格均衡似乎就是唯一稳定的均衡。同理，我们不必过分的强调此结果的意义，因为，就像我们后面所证明的那样，只要稍稍详细地考察模型，就会发现单一价格均衡可能是不存在的。

---

24 为简单起见，假设每个企业只拥有一台机器（一个职位），由于每单位时间搜寻数目为 s，所以每单位时间向该企业递交申请的人数为 sL/N，其中 L 是工人数目，而 N 是职位数目。于是支付工资水平为 w 的企业每单位时间所能够填满职位空缺的可能性为 $\gamma = (sL/N)F$。通过简单地计算可以得到零利润条件：

$$K = \frac{\gamma}{\gamma + r}\frac{P}{r + q} = 1 - \frac{\gamma}{\gamma + r}\frac{q}{r + q} \tag{17}$$

其中 $q = s(1-F) + \mu$ 而 $P = a - w - (q + r)T$，式（16）、式（17）的解法同以前类似，如此可以求得均衡工资分布。而每个企业的利润都相等。L/N 的值保证了利润水平为零。

25 前面我们证明了如果条件式（7）不满足，那么就不存在零利润均衡：企业总是会把名义工资水平降低到可以接受的最低水平上，从而利用了由于搜寻成本产生的垄断力量。如果条件式（7）满足，那么就存在零利润的多重均衡以及正利润均衡，如果条件式（7）不满足，那么就只是存在正利润均衡（准租金），而且其形式就是我们在前面部分所描述的那样。

26 如果经济中有多于一种商品，那么即使条件式（7）不满足，也可能存在着充分就业的利润的均衡。假设还存在着一个产业部门不需要专门培训，令此部门的产出为计价单位，由于企业进入原来的部门需要培训，相对于正利润的情况，价格下降，利润也下降了。有关两部门模型更加详细的阐述，见“Equilibrium Wage Distribution” IMSSS，1974（中文译文见本卷“均衡工资分布”）。

此处要求培训成本不仅仅由有培训成本的部门承担。

三、模型的稳健性

在第二部分构造模型的过程中，我们引入了一系列简单的假设。我们已经考察了放宽这些假设的结果。根据上述详细的分析，我们肯定此模型是稳健的。即便可以，现在就推导出均衡工资分布以及非自愿失业条件下的均衡情况还为时尚早。在这个部分，我们将简要考察模型的两种扩展情形：企业能够进行技术选择和工人能够进行搜寻强度的选择（区别于技术选择条件下的假设）。在下一个部分，我们将会考察一个重要扩展，企业存在非独有特征区别的情况。

1. 技术选择

假设企业能够选择技术并不会对分析结果产生重要的影响。但是它的确会增加我们能够得到的工资分布类型。假设单位工人产出函数是培训成本的递增的凹函数

$$a'(T) > 0, a''(T) < 0 \tag{18}$$

双重工资水平的均衡分布。在双重工资水平条件下，零利润均衡条件可以由零利润方程得到：

$$a(T_1) = w_1 + (\mu + r)T_1, a(T_0) = w_0 + (\mu + r + s\pi)T_0 \tag{19}$$

最优的技术选择方程为

$$a'(T_1) = \mu + r, a'(T_0) = \mu + r + s\pi \tag{20}$$

而下列不等式确保了选择了技术 T_1 的企业会支付 w_1 的工资水平

$$\frac{w_1 - w_0}{s\pi} < T_1, \frac{w_1 - w_0}{s\pi} > T_0 \tag{21}$$

图5用图形将此结论显示出来。由式（20）可以解出选择了较高培训成本的技术的企业工资水平：在所有均衡条件下都是一样的。这样通过式（19）就可以解得在选择该技术条件下获得零利润的工资水平。考虑一种工资水平 w_0。过点 w_0 画出与曲线 a（T）相切的直线，令支付较高工资水平的企业比例为 $\mu + r + s\pi$ 这正好是这条线的斜率。那么由式（20）可以解出支付了 w_0 的工资水平的企业会选择技术 T_0，并且通过式（19）就可以知道该企业是盈亏相抵的。

注意到在我们所构造的这个例子之中，培训成本的差异是模型内生的。它并不是外生的企业特征。

而且我们可以类似地构造出任意有限个工资水平的均衡。

连续分布的工资水平。如果工资水平是服从连续分布的话，那么零利润

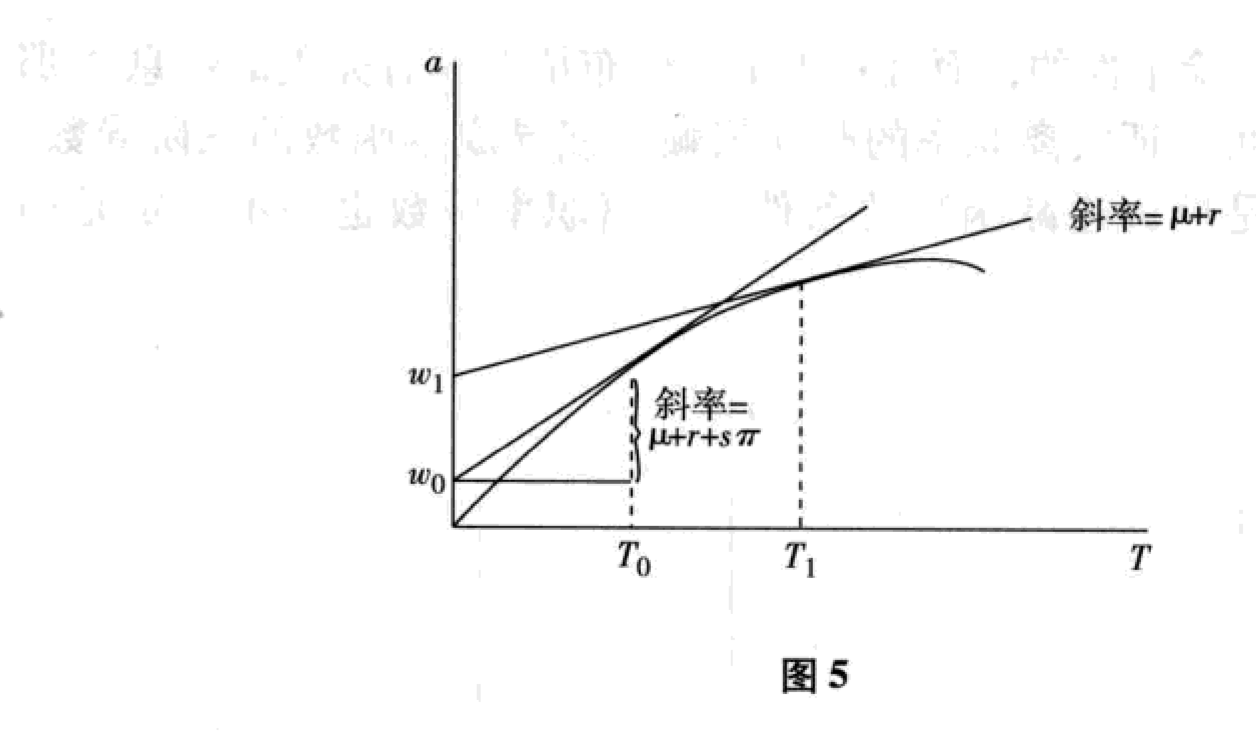


图 5

条件意味着对于所有观察到的 $\{w, T\}$，有

$$a(T) = w + T\{\mu + r + s[1 - F(w)]\} \tag{22}$$

对于 w 和 T，求一阶条件得到

$$a'(T) = \mu + r + s[1 - F(w)] \tag{23}$$

和

$$1 - sfT = 0 \tag{24}$$

将其代入得到

$$a'\left(\frac{1}{sf}\right) = \bar{\mu} + r + s(1 - F) \tag{25}$$

这样工资水平分布密度函数就仅仅依赖于产出函数 a（T）。

2. 搜寻成本和搜寻强度的决定

如果搜寻行为会产生相关的成本，那么作为理性的工人，在决定搜寻次数的时候，会比较各个搜寻强度与搜寻成本。假设搜寻成本是搜寻强度的递增凸函数是很合乎情理的，也就是说，如果，C（s）是在搜寻强度为 s 的条件下单位时间的搜寻成本，有

$$C' \geqslant 0, C'' \geqslant 0 \tag{26}$$

在前面的分析之中，我们假设成本函数的具体形状如图 6（b）所示，这就意味着均衡条件下是不存在搜寻行为的。

正确计算出搜寻的收益不是很容易的。我们必须知道每个工人对于工资水平分布信息的了解程度，他们的风险厌恶程度，以及他们预期能够维持某个职位的时间长短。例如，倘若工人不知道工资分布情况，那么搜寻不仅仅可以直接提高找到职位的可能性，而且还会间接地增加工人对于工资分布情况的了解，从而可以在相同的搜寻强度条件下做出更好的决策。为简单起

见，假设工人是风险中性的，而且对于工资分布情况具有完美的信息。[27]那么简单的计算就可以证实离职率的凸性依赖于搜寻成本函数的三阶导数，显然，即使在f是单调递减函数的条件下，离职率函数也并不一定是凸的。

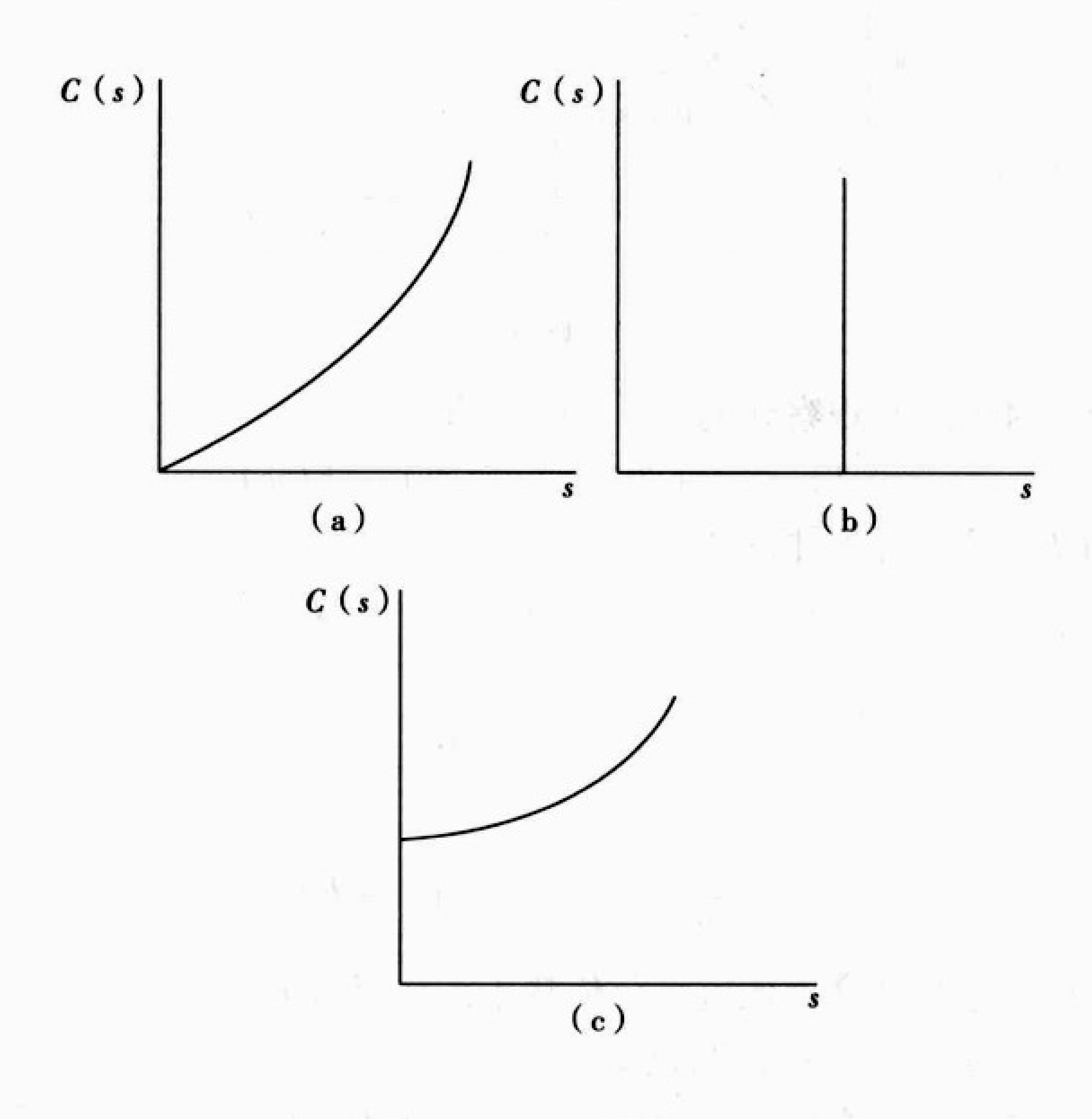


图 6

27　工资水平为w的工人其期望收入的净现值为

$$V(w)=\max_{s}\frac{w-C(s)+s\int_{w}^{w^{+}}V(\tilde{w})dF(\tilde{w})}{\mu+r+s[1-F(w)]}\tag{27}$$

因此s的选择必须满足

$$C'(s)=\int_{w}^{w^{+}}V(\tilde{w})dF(\tilde{w})-[1-F(w)]V(w)\tag{28}$$

其中二阶条件要求$C''>0$。由式可以得到搜寻强度作为支付工资水平的函数（给定分布函数$F(w)$）。

我们对第一部分的扩展方式是非常简单明了的。[28]其中的一个变化在于，在我们前面的福利分析之中，唯一可能存在的无效率源于培训成本的过量支出，在内生搜寻条件下，无效率的情况还可能是由于搜寻成本的过量支出造成的。

一个替代方案。此模型的一个替代方案是由迪布维格（Dybvig）和杰恩斯（Jaynes）提出来的。每个人对每一个职位都递交申请，但是只接受工资最高的职位。企业在求职者之中随机地选择。于是，支付工资水平为 w 的企业将会有 $[\mu+F(w)]L$ 的求职意向。其中 L 是总的劳动力水平。在稳态均衡之中，支付工资水平为 w 的企业将会有 $\mu Lf(w)$ 的求职意向。因此，工人从工资水平为 $\hat{w}$ 的企业跳槽去工资水平为 w 的企业的可能性为 $\mu f/(\mu+F)$。[29]因而总的离职率为

$$q(\hat{w}) = \int_{w}^{\infty} \frac{\mu f(w)}{\mu + F(w)} dw$$

其分析过程与第一部分完全一样。[30]

四、非金钱回报

职位的差异不仅仅反映在工资水平上，还反映在其他一些非金钱层面。

---

28 有一种情况的确会产生这个问题：如果搜寻成本形状如图 6（c），那么对于支付工资水平最高的企业 $s=0$，离职率函数在 $w=w^{+}$ 处为水平的，因此在存在搜寻成本的条件下，没有企业愿意支付 w^{+} 的工资水平。相应地，如果均衡之中存在着工资差异，那么就会产生所有企业支付的工资水平都为最低工资水平的退化的差异工资分布的情况，但是即便如此，也不可能产生零利润均衡。当所有企业都支付 $w=w^{-}$ 时，确实存在一个非零利润均衡。此结果同 Diamond（1971）的结果一致，在他的搜寻模型之中，他证明了均衡包括一个单一的垄断价格水平均衡。尽管图 6 给出的成本函数可能有一点不合情理，但是即使在多重工资水平条件下均衡也同样可能存在，条件是工人决策时将企业非独有特征的回报纳入考虑范围。或者工人们之间的搜寻成本是有差异的，其中一部分工人的成本满足 $C'(0)=0$，或者存在着别的信息技术（比如工人可以通过买报纸得到当时所有的工资水平，Salop and Stiglitz，1977）。

29 这里我们忽视了在某个时期工人同时获得两个职位的可能性。如果时期是很短的，因而相对于求职者人数而言，空缺的职位数是很小的，这种情况就是小概率事件；在我们连续时间假设条件下，这种情况就是零概率事件。

30 这里有双重工资分布（参见图 7），其中 $w_1=w^{*}$，$w_0=w_1-[\pi\mu T/(1-\pi+\mu)]$。因为所有的低工资工人以及新加入劳动力的工人都会应聘 $\mu\pi L$ 的高薪企业的职位，他们得到更高薪水职位的概率为 $\mu\pi/(1-\pi+\mu)$。对任意的组合（w_0，π），只要满足 $w_0=w_1-[\pi\mu T/(1-\pi+\mu)]$ 都是均衡的。对于连续统一 $-q'T=1$，或者 $(\mu f/\mu+F)T=1$ 即 $F=ke^{w/\mu T}-\mu$ 对于所有 $w^{-}\equiv\mu T\ln\mu/k\leqslant w_0\leqslant w\leqslant w^{-}\equiv\mu T\ln(1+\mu)/k$，$k<\mu$ 均成立。如果 $w_0>w^{-}$，那么在点 $w=w_0$ 处存在一个聚点；而在点 w^{+} 处永远不存在聚点。

某些个人可能会喜欢某个企业的某些特征，而其他的工人可能喜欢另外一个企业的某些特征。为简单起见，假设工人在完成培训从而完全了解该企业之前，是不知道这些特征的。[31]我们将非企业独有特征的回报的消费等价值记为 θ，并且不失一般性令 $E\theta=0$。[32]我们将非独有特征的回报的分布函数记为 $M(\theta)$，即在向某个企业求职的工人是随机样本条件下，$M(\theta)$ 能够体现他们对于非企业独有特征的评价值小于或者等于 θ。我们进一步假设工人对于企业 i 的评价与对于企业 j 的评价是相互独立的。而且为简单起见，假设对于所有企业它的分布是相同的。我们将密度函数记为 $m(\theta)$。工资水平为 w，并且认定某企业非金钱特征为 θ 的风险中立的工人只有在工作水平 $\hat{w}$ 满足下列条件才会接受该职位：

$$\hat{w} \geqslant w + \theta - \tilde{\theta} = w_r \tag{29}$$

对它而言，$\hat{w}$ 是保留工资水平。[33]因此，离职率函数就是

$$q(w) = \mu + \int s[1 - F(w + \theta - \tilde{\theta})]\tilde{m}(\theta, w)d\theta \tag{30}$$

其中是评价为 θ，工资水平为 w 的企业分布函数 $\tilde{m}(\theta, w)$（可能会与 $M(\theta)$和 $F(w)$ 相关）因此，

$$q' = -\int sf(w + \theta)\tilde{m}(\theta, w)d\theta + \int s(1 - F)\tilde{m}_w(\theta, w)d\theta \tag{31}$$

式（31）的重要意义在于，即便是工资水平最高的企业也能够通过提高工资水平来降低跳槽率。如果该企业试图降低工资水平的话，就算存在着搜寻

31　当然，某些非金钱特征在接受职位之前是能够知道的；这会对企业的聘用率产生影响。同样，某些非金钱特征，比如提升率只有在接受职位之后才能够知道的。

32　如果工人认为某个企业的非独有特征要优于另一个企业的话，那么我们只要把 θ 的均值加入到工资水平里面就可以了。

33　为简单起见，将第一部分搜寻模型的分析范围局限于 s_1 是固定的情况。对于更加一般的情况，见"Equilibrium Wage Distribution"，IMSSS，1974。$\tilde{\theta}$ 表示某个在收到工资水平为 $w=w_r$ 职位之后对于是否再进行搜寻是无差异的。也就是说，如果不进行搜寻，他的效用水平 $w+\tilde{\theta}$ 等于搜寻的效用水平。在单一工资水平的特例之中，保留评价水平为 $\theta=\tilde{\theta}$ 的工作期望效用值（净现值）为 $w^* + \tilde{\theta}/\mu + r$。如果他选择跳槽，那么他找到工作还需要一些时间。如果他真的选择跳槽，θ 可能会小于 $\tilde{\theta}$，这样的话他又会跳槽；如果 θ 大于 $\tilde{\theta}$，这样他就会保留这份工作。因此，他的期望效用值（净现值）为 $\frac{\mu + s(1 - M(\tilde{\theta}))}{\mu + r}\int_0^{\infty}\exp\{-rt\}\left\{w^* + \int_{\tilde{\theta}}^{\infty}\frac{\theta dM(\theta)}{1 - M(\tilde{\theta})}\right\}\exp\{-[s(1 - M(\theta)) + \mu]t\}dt\,\frac{w^*}{(\mu + r)}$ $\frac{[s(1 - M(\tilde{\theta})) + \mu]}{[r + s(1 - M(\tilde{\theta})) + \mu]} + \frac{s\int_{\tilde{\theta}}^{\infty}\theta dM(\theta)}{(\mu + r)(r + s(1 - M(\theta)) + \mu)}$，于是通过解下列方程得到 $\tilde{\theta}$ 值 $\mu - rw^* = s\int_{\tilde{\theta}}^{\infty}\theta dM(\theta) - \tilde{\theta}(r + s(1 - M(\theta)) + \mu)$。

成本，该企业的跳槽率也将会增加。此均衡有着一些十分有趣的特性。

1. 不可能存在单一工资水平均衡条件下实现充分就业

如果所有企业都支付相同的工资水平，那么所有发现 $\theta < \tilde{\theta}$ 的工人都会搜寻更好的职位。而且在接受企业职位并且参加完培训之前，工人也不知道该职位的特征，因此他们只能接受第一个职位（因为所有企业都支付相同的工资水平）。

如此就可以很容易地计算出在支付工资水平 $w > w^*$ 时企业的离职率（在稳态情况下）：所有满足 $\theta < w^* - w + \tilde{\theta}$ 的工人都会选择离职。

那些没有选择离职的工人构成了企业的劳动力的主要成分。其中评价值满足 $\theta < w^* - w + \tilde{\theta}$ 的工人的平均在职时间为 $1/(\mu + s)$。而评价值满足 $\theta > w^* - w + \tilde{\theta}$ 的工人的平均在职时间为 $1/\mu$，因此离职率就是

$$q(w,w^*) = \frac{1}{\dfrac{M(w^* - w + \tilde{\theta})}{\mu + s} + \dfrac{1 - M(w^* - w + \tilde{\theta})}{\mu}} \tag{32}$$

其中

$$\frac{\partial q}{\partial w} = \frac{-q^2 sm(w^* - w + \tilde{\theta})}{\mu(\mu + s)} \tag{33}$$

由于存在着单一价格均衡

$$\left.\frac{-\partial q(w,w^*)}{\partial w}\right|_{w=w^*} = \frac{sm(\tilde{\theta})\mu(\mu + s)}{\{\mu + s[1 - M(\tilde{\theta})]\}^2} = \frac{1}{T} \tag{34}$$

但是，这并不是最一般情况。[34] 所节省下来的跳槽成本并不足以维持单一工资水平。

2. 存在非自愿失业的均衡

如果

---

34 尽管在前面的例子里，最优的工资水平为常数。但是这里的情况是不同的。这是因为那些在本企业在职时间较长的工人们披露了他们对于企业非金钱特征的评价信息，即 $\theta > \tilde{\theta}$，该企业就可以利用这个信息来设计出一套资历工资结构。这个问题是很复杂的，但是其结果不会改变本部分所推出结果的性质。同理，可以推出，式（31）表示的一阶条件并不完全准确。如果企业选择采取固定工资率，那么任何工人的离职的概率都是在职时间长短的函数，即 $q = q(w, t)$，企业利润的净现值为

$$P = \int q(a - w)\exp\{-[r + \bar{q}(w,t)]t\}dt - T$$

其中 $\bar{q}(w,t) = \dfrac{\int_0^t q(w,t)dt}{t}$ 在我们的分析过程之中，我们将 $\bar{q}(w, t)$ 替换为 $\bar{q}(w, \infty)$。同样，对于所有企业如果满足 $w \geq w^-$ 并且 $w = w^*$，都有 $\dfrac{\partial P}{\partial w} \neq 0$ 这个核心结论仍然是成立的。

$$\left.\frac{-\partial q(w,w^*)}{\partial w}\right|_{w=w^*} = \frac{sm(\tilde{\theta})\mu(\mu+s)}{\{\mu+s[1-M(\tilde{\theta})]\}^2} > \frac{1}{T} \tag{35}$$

那么均衡就一定存在着非自愿失业。[35]失业率越高，降低工资水平所造成的离职率上升的边际效应就越小。

令 U 为失业率；很容易看出，一般而言企业的离职率都是依赖于失业率以及所有企业支付的工资水平 w^*。[36]

$$q = q(w,w^*,U)$$

于是均衡失业率就是下列方程组的解：

$$-Tq_w(w^*,w^*,U) = 1 \tag{37}$$

$$a = w^* + [q(w^*,w^*,U) + r]T \tag{38}$$

3. 存在着多重工资水平的均衡

此时，人们很自然地要问，是否会存在着多重工资水平的零利润的均衡；换句话说，我们的模型之所以存在着难题是否源于对所有企业都必须支付相同工资水平的限制？答案是否定的：一般情况下，不存在多重工资水平条件下的零利润的均衡。[37]

4. 摩擦性失业

如果对于现在的模型进行稍稍的修改，我们就能够得到另一种完全不同类型的失业。假设只有失业者才能够进行搜寻，或者更加一般地说，在失业状态下进行搜寻的成本相对较小，那么均衡就能够由一定的摩擦失业率来表示。凡是觉得就职企业的非独有特征的评价值很小的工人都会跳槽，他们宁愿保持失业状态直到找到新的职位空缺为止。

5. 自然失业率的非优性

此处的均衡失业率可能会被视为自然失业率。然而，这种失业率的福利

---

35 如果对于所有

$$w \geqslant w^-，有\left.\frac{\partial q(w,w^*)}{\partial w}\right|_{w=w^*} < \frac{1}{T}, \tag{36}$$

式（36）那么就存在着正利润均衡，表明企业希望降低工资水平直到工人能够接受的最低工资水平 w^-。一般而言，在 w^- 处，有 $a > w + (q+r)T$，也就是说即存在着正利润。此分析过程与第二部分的方法是一致的。

36 由下面的搜寻和非金钱特征假设条件推导出离职率函数是非常烦琐的；对于结果重要的是当 $w = w^*$ 时，$q_wU(w, w^*, U) > 0$。

37 增加差异工资水平的数目并不能够解决这个难题，因为对于每一个工资水平而言，有两样信息（工资率和支付该工资率的企业比例），以及两个方程（零利润方程和利润最大化方程）我们都不知道。正式的证明见本文 NBER 工作论文版（1984）。

性质并不明显。福利分析过程类似于第二部分第1点讨论两种附加效应的方法。[38]首先，每一个企业都相信，如果它们提高相对工资水平，就能够降低跳槽率，当然了，如果它们都选择提高工资水平的策略，那么结果只是提高了整体工资水平，降低了劳动力的需求，而不会对跳槽率产生直接的影响；其次，劳动力需求的降低，也就是失业率的上升，会产生严重的外部性：它降低了所有企业的跳槽率，这样会降低整体的培训成本，使得工资率最后会增加。尽管企业进行决策的时候不会考虑失业率上升所带来的好处，它同样不会考虑这样所带来的成本：企业职位与工人不匹配，年轻工人的推迟就业，以及由于企业职位与自身不匹配而跳槽的工人推迟就业。

五、结论

本文考察了不完美信息条件下工资水平均衡分布情况。我们证明了即使是在稳态均衡条件下，同质的劳动力是如何获得差异工资水平的。原来我们认为会在长期消除工资水平差异的竞争性力量，在不完美信息条件下是无法起作用的。实际上，即使不完美信息仅仅是源于市场上总是存在的工资水平差异，情况也是如此。只要市场上存在着由于企业（对称的）非金钱特征以及工人偏好差异所产生的不完美信息，市场上就一定会存在工资水平的差异，而多重工资水平的零利润的均衡就不会存在。

也许本文最重要的结果就是在均衡条件下竟然会存在非自愿失业——尽管存在着劳动力的超额供给，仍然没有企业愿意在更低的工资水平条件下雇佣工人。这是因为企业知道，一旦它们这么做，离职率就会上升，跳槽成本（培训成本）就会增加，最终利润水平实际上会降低。因此，这个模型为名义或者实际工资刚性提供了一个基本解释。另外，此模型还提出了一个摩擦失业理论：当工人和企业的职位不匹配时（就企业的非金钱特征而言），他们会回到失业大军之中谋求更好的职位。

尽管通过有限最优性（将获取信息以及搜寻的成本视为外生）可以引致非自愿失业以及工资水平差异，但是一般而言，失业率以及工资差异水平在均衡条件下不是最优的。不幸的是，政府干预以实现有效的帕累托改进的必要性，远远不是一眼就可以看出来的。

38　正式的比较最优单一工资水平均衡以及单一工资条件下的市场均衡，见本文的 NBER 工作论文版（1984）。

参考文献

Akerlof, G. (1984). 'Gift exchange and efficiency, wage theory: four views'. *American Economic Review Proceedings*, Vol. 74, no. 2, pp. 79 – 84.

Arnott, R. and Stiglitz, J. E. (1985). 'Labor turnover, wage structures, and moral hazard: the ineffciency of competitive markets'. *Journal of Labor Economics* (forthcoming).

Butters, G. R. (1977), 'Equilibrium distributions of sales and advertising prices,' *Review of Economic Studies*, vol. 44, pp. 465 – 91.

Calvo, G. (1979). 'Quasi-Walrasian theories of unemployment,' *American Economic Review*, vol. 69, pp. 102 – 8.

Crawford, V. P. (1983). 'Long-term relationships governed by short-term contracts.' International Centre for Economics and Related Disciplines, London School of Economics.

Diamond, P. A. (1971). 'A model of price adjustment.' *Journal of Economic Theory* vol. 3, pp. 156 – 68.

Dybvig, P. and Jaynes, G. (1980). 'Output-supply, employment and intra-industry wage dispersion.' Cowles Foundation Discussion Paper no. 546.

Grossman, S. and Stiglitz, J. E. (1976). 'On the impossibility of informationally efficient markets.' *American Economic Review*, vol. 70, pp. 393 – 408.

Grout, P. A. (1984). 'Investment and wages in the absence of binding contracts: a Nash bargaining approach,' *Econometrica*, March, pp. 449 – 60.

Leibenstein, H. (1957). *Economic Backwardness and Economic Growth*, New York: Wiley.

Mortensen, D. T. (1970). 'A theory of wage and employment dynamics.' In Microeconomic Foundations of Inflation and Employment Theory (ed. E. Phelps et al.). New York: Norton.

—— (1973). 'Search equilibrium in a simple multi-market economy.' Center for Mathematical Studies, Discussion Paper no. 54, Northwestern University.

Nalebuff, B. J. and Stiglitz, J. E. 'Quality and Prices.' Princeton Mimeo.

Reinganum, J. F. (1979). 'A simple model of equilibrium price dispersion.' *Journal of Political Economy*, Vol. 87, pp. 851 – 8.

Salop, S. C. (1973a). 'Systematic job search and unemployment.' *Review of Economic Studies*, vol. 40, pp. 191 – 201.

—— (1973b). 'Wage differentials in a dynamic theory of the firm.' *Journal of Economic Theory*, vol. 5, pp. 321 – 44.

—— (1977). 'The noisy monopolist: imperfect information, price dispersion and price discrimination.' *Review of Economic Studies*, vol. 44, pp. 383 – 406.

——and Stiglitz, J. E. (1977). 'Bargains and ripoffs: a model of monopolistically competitive price dispersion.' *Review of Economic Studies*, vol. 44, pp. 493 – 510.

——and—— (1982). 'The theory of sales: a simple model of equilibrium price dispersion with identical agents.' *American Economic Review*, vol. 72, no. 5, pp. 1121 – 30.

Shapiro, C. and Stiglitz, J. E. (1984). 'Equilibrium unemployment as a worker discipline device.' *American Economic Review*, vol. 74, no. 3, pp. 433 – 44.

Stiglitz, J. E. (1976). 'Prices and queues as screening devices in competitive markets.'

IMSSS Technical Report no. 212. Stanford University.

—— (1982). 'Alternative theories of wage determination and unemployment: the efficiency wage model.' *In The Theory and Experience of Economic Development: Essays in Honor of Sir W. Arthur Lewis* (eds. M. Gersovitz, C. F. Diaz-Alejandro, G. Ranis and M. Rosenzweig), pp. 78 – 106. London: George Allen & Unwin.

—— (1984). 'Theories of wage rigidity.' NBER Working Paper no. 1442.

Weiss, A. and stiglitz, J. E. (1980). 'Job queues and lay-offs in labour markets with flexible wages', *Journal of Political Economy*, vol. 88, pp. 526 – 38.

Yellen J. (1984). 'Efficiency wage models of unemployment.' *American Economic Review Proceedings* vol. 74, no. 2, pp. 200 – 6.

销售理论：同质代理人条件下的价格离差均衡模型*

新古典经济学理论的一个基本原则就是一价定律。关于单一价格是如何决定的问题，我们可以从有关瓦尔拉斯拍卖人的寓言故事中得到一些启发。本文的目的正是要检验在传说中的（瓦尔拉斯）拍卖人缺位以及在信息收集成本的条件下，完全竞争市场中均衡的问题。在这种条件下，个人对于出售商品的价格或质量可能并不完全了解。与新古典（经济学）理论传统所研究的情况相比，该条件下的均衡可能会有显著不同。特别地，即使对于同质商品，唯一的市场均衡也可能存在价格离差；于是一价定律就不再成立了。我们将用一个模型来加以阐释。在这个模型中，所有个体都是同质的（事前），没有外生的扰动源，不存在市场外部的扰动，因此市场总是出清的。相反，扰动完全是由市场内部的运行引起的。所以，这里的信息不完美是由市场自身造成的。

在我们的模型中，虽然（假定）所有消费者都具有相同的偏好和收入，并且所有厂商都采取同样的技术，[1] 依然会存在一些定价较高的厂商和一些定价较低的厂商。那些（很不走运地）碰到较高价格厂商的消费者仅仅购买当期的需求量，然后在下一期再进入市场（购买）；那些（很走运地）碰到较低价格厂商的消费者会“经济”地多买一些，以供将来消费。尽管价格高的厂商的每一笔交易的利润要高一些，但是，其交易量要低一些，均衡条件要求两种厂商的利润相等，就是说（高价厂商的）低销售量完全从较高的价格得到补偿。

---

* S. Salop and J. E. Stiglitz, “The Theory of Sales: A Simple Model of Equilibrium Price Dispersion with Identical Agents, *The American Econonomic Review*, Vol. 72, No. 5 (Dec., 1982), 1121 - 1130. 摩根斯特恩（*Oskar. Morgenstern*）高等奖学金并且作为普林斯顿大学高级研究中心的客座教授期间下笔的。同时非常感谢国家科学基金会给予的资助。本文最早刊于1978年2月贝尔实验室关于垄断竞争与不完美信息会议文集。R. Lindsey, V. Krishna 以及 D. Abreu 对本文的评论令我们获益良多。

1 假设同质的消费者和厂商以及没有外生扰动是为了向大家清楚地证明，这里所分析的价格离差明显不同于以前的模型里所分析的情况（比如，Salop，1977；Salop and Stiglitz，1977），在那里存在价格离差是为了区分出不同的消费群。

我们建立的模型是非常有趣的，不仅因为它启发我们关注经济中价格离差（现象）的本质；还因为它就零售（业）的某些方面，给出了至少部分的解释，而这些问题本来是很难解释的。

关于第一部分里所采用的基本模型有一个简化的假设，那就是没有市场进入成本。一旦存在进入市场的成本，那么唯一可能的（多重）均衡必将包含价格离差；而在不存在均衡价格分布的条件下，市场均衡根本就不存在。因为当不存在扰动时，厂商会不断提高价格，直到消费者剩余为零，当然前提是消费者已经进入了市场。而低价厂商所提供的（预期）消费者剩余必须足够大，使得消费者愿意支付与再次进入市场相关的所有固定成本。这样，存在价格离差的一个必要条件就是厂商无法对（不同的）消费者实行完全的价格歧视。原因在于，如果厂商有能力采取某些措施（比如非线性定价策略）对消费者的不同人群（这里分为老年人和年轻人）实施价格歧视的话，就不会存在价格离差的均衡了。而且，一旦存在市场进入成本，也不会有均衡存在。

尽管本文中所采用的模型是非常简单的，但是，它的基本定性结果是十分可靠的，在本文的第二部分，我们注意到了模型有待扩展的一系列方向，比如，在本文（以及大多数相关文献）中，都假设了理性预期的存在：消费者拥有的关于价格分布的信念会影响到他的决策——到底是现在就购买未来的需求，还是等到下一期再进入市场；并且信念与均衡中的（真实）价格分布相吻合。我们会证明，即使在理性预期不成立条件下，仍然会存在（均衡）价格分布，实际上，与理性预期条件相比，在有些条件下更可能出现（均衡）价格分布。

一、销售的理论

尽管下文中的价格离差模型是高度简化了的，但它的确抓住了这类经济现象的本质：价格离差可能发生在不同厂商之间、同种商品的不同品牌之间或者是同一厂商的不同时期之间；具有离差的变量可以是价格，也可以是质量。有一个特别有意思的现象，罗伯特·斯特纳（Robert Steiner，1978）发现：大多数进行全国性广告宣传的商品，不同商店的销售价格在同一时间里是相同的，但是每个商店都会不时地打折，或者是向消费者让渡厂家折扣，或者将有广告的商品连同没有广告的商品进行打折处理。下面的模型只

考虑没有广告的特价商品。[2]

我们假设每个消费者存在两期，当商品价格不高于某一个保留价格（reservation price）u 时，每个消费者在每一期都有一单位的商品需求，因此，一个垄断厂商会选择价格 $p^m = u$，我们把这个价格称为“垄断价格”。在出现价格离差的情况下，进入市场的消费者或者在两期各买一单位商品或者在第一期购买两个单位商品，消费掉一个单位，把剩余商品存储起来以供第二期消费。倘若消费者选择为了储备而购买，就节省了由于再次进入市场产生的额外的交易成本 c。但是，这样就会产生存储成本 δ，[3]到底是购买后再储备，还是下期购买取决于如何平衡考虑这两种成本。

假设消费者先验地知道一个关于市场上某种商品价格的分布函数 $f(p)$，在缺乏更具体信息的条件下，消费者会在第一期随机地选择一家厂商，假定这家厂商价格为 p，令 $\hat{p}$ 为消费者的保留价格，在这个价格下，对于本期是为了储备购买，还是只购买当期需求量，等到下期再进入市场（购买），消费者是无差异的。为了重点考察不同时期间的交易成本，我们假定消费者不能拒绝报价 p 而在第一期再选择一家厂商，也就是说，我们假设在任何一期，找寻第二家厂商的成本是如此之高，以至于消费者永远不会这么做。[4]

注意到下一期消费者会面临平均价格 $\bar{p}$，而且还要支付交易费用 c，$\hat{p}$ 由下式给出[5]

$$\hat{p} + \delta = \bar{p} + c \tag{1}$$

1. 均衡的描述

我们希望均衡价格分布函数 $f^*(p)$ 具有以下特性：

---

2 必须要强调的是，我们的模型是高度理想化的，这是为了阐明在信息不完美的条件下零售行为的某个特征。特别是，消费者除了亲自去商店以外，还可以通过各种渠道来获取信息。比如可以通过个人视察、购买体验、销售商（包括广告）和卖信息的人等来获得。

3 在本文我们假设存储成本与所购买的单位数量成正比，而不依赖于购买价格。如果存储成本包含利率和损耗，那么它就和价格成比例。但这样假设不会对前面的分析有实质性的改变。

4 这里不同于大部分关于搜寻的文献。那些文献通常假设每次的搜寻成本为常数。我们的假设简化了分析，但确不是关键性的。依照关于搜寻文献的惯例，我们假设消费者知道价格的概率分布。由于先验的概率分布对每一个厂商都一样，因此，到每个厂商光顾的消费者数量也是一样的。

5 这就假设了消费者是风险中性的。在我们 1976 年合作的第二篇文章里，我们证明了这个假设不是关键性的。另外，需要有

$$\hat{p} \leqslant u - \delta \text{ 成立，（否则，消费者购买储量就不划算了）} \tag{1'}$$

在更一般的情况下，写作

$$\hat{p} = \min\{\bar{p} + c + \delta,\ u - \delta\} \tag{1''}$$

1）利润最大化：在给定其他厂商出价的条件下，每个小厂商都选择价格 p，以便利润最大化；

2）等利润：每个厂商都得到相同的（正常）利润；

3）搜寻的均衡：每个消费者在 $f^*(p)$ 给定的情况下寻求最优（消费路径）；

4）进入的均衡：当且仅当预期的消费者剩余为非负的情况下，消费者才会进入市场。

我们会推导出非退化的均衡价格分布存在所必需的条件。为此，我们必须先说明出这种均衡（如果它存在的话）的特点。我们证明了这个均衡中只能有两个价格，较高的一个等于垄断价格，而较低的一个等于保留价格，然后，我们推导出了在垄断价格和保留价格两个价格条件下利润相等的条件。

首先证明

引理1：在均衡价格分布中，至多存在两个价格。

假设存在不止一种低于 $\hat{p}$ 的价格，比如 p_m 和 p_l，并且 $p_m > p_l$，价格为 p_m 和 p_l 的厂商将面临相同数量的消费者，这些消费者将都购买两期的量，这里价格 p_m 厂商的利润必然大于价格 p_l 的厂商。于是无法得到等利润的均衡。用类似的逻辑，我们可以除去两种价格都比 p_m 要大的情形。

2. 双价格均衡

记这两种价格分别为 p_l 和 p_h，而价格为 p_l 和 p_h 厂商所占份额分别为 $(1-\lambda)$ 和 λ。

引理2：任何双价格的均衡都必须具有如下性质：低价格厂商索取的价格就是消费者的保留价格：

$$p_l = \hat{p} \tag{2}$$

如果 $p_l > \hat{p}$，消费者就会先后向两家厂商各买一期的数量，但是由于 $p_h > p_l$，两种价格所产生的利润不可能相等，如果 $p_l < \hat{p}$，消费者会向低价格的厂商一次性购买两期的需求量。但是低价格的厂商即使稍稍提价也不会损失任何客户，而且还可以增加利润，因此，p_l 不可能是利润最大化的均衡价格。

引理3：在任何双价格均衡中，高的价格必然是垄断价格：

$$p_h = u \tag{3}$$

假设不是这样，由于 $\hat{p} < p_h$，价格为 p_h 厂商只能卖出一期的量，要是他准备提高价格，他将不会损失任何客户，于是他的利润会上升。

利用等利润条件可以计算出价格为 p_h 的厂商的份额 λ，过程如下：假设共有 L 个消费者和 n 个厂商，每个厂商可以吸引来 L/n 个年轻消费者，以及 $\lambda L/n$ 个老年消费者，这些老年消费者在年轻时很不走运地碰见了一个高价厂商，因此只买了一期的量。[6]由于不走运的年轻人和老年人都只在每期购买一个单位，每个价格为 p_h 的厂商的销售量 X_h 可以给出：

$$X_h = (1 + \lambda)L/n \tag{4}$$

索取价格 p_l 的厂商的销售量 X_l 要高一些，每个厂商都售出 2 个单位给年轻的消费者，1 单位给老年消费者，即

$$X_l = (2 + \lambda)L/n \tag{5}$$

为简单起见，我们假设生产的边际成本为零，于是等利润条件意味着

$$p_h X_h = p_l X_l \tag{6}$$

然后利用式（3）、式（4）和式（5），可以得到

$$(1 + \lambda)u = p_l(2 + \lambda) \tag{7}$$

代入 $\bar{p}$ 的定义

$$\bar{p} = \lambda p_h + (1 - \lambda)p_l \tag{8}$$

注意到保留价格式（1），再利用式（2），我们得到了：

$$p_l = \lambda p_h + (1 - \lambda)p_l - (\delta - c) = u - (\delta - c)/\lambda \tag{9}$$

式（7）和式（9）给出了包含两个未知数 p_l 和 λ 的两个方程，这样，就可以求出市场均衡解，由此我们建立：

定理 1：如果存在一个非退化的价格离差均衡，它必然包含两个价格：[7]

$$p_l = (u + (\delta - c))/2, p_h = u \tag{10}$$

并且要高价的厂商所占总数的比例为

$$\lambda = 2(\delta - c)/(u - (\delta - c)) \tag{11}$$

为了忽略掉消费者是否进入市场的问题，我们着重考察当 $c = 0$ 的特别情形，更一般的情况会在第二部分讨论，现在我们可以建立：

定理 2：如果 $c = 0$，包含双价格的均衡的必要和充分条件是：

$$\delta < u/3. \tag{12}$$

为了让式（10）和式（11）构成有双价格的均衡，必须满足 $0 < \lambda < 1$

---

6 假设他们都在第二期再次进入市场。参见我们 1981 年的文章的第三部分，以及本文下文。

7 注意包含双价格均衡（Two Price Equilibrium，TPE）的性质，如果存在的话，它仅仅取决于第二期的搜寻成本，因此，如果 c_i 是第 i 期进入市场的成本，则有

$$p_l = \frac{u + (\delta - c_2)}{2}，\lambda = \frac{2(\delta - c_2)}{u - (\delta - c_2)}$$

以及 $p_l+\delta \leqslant u$。第二个约束确保了消费者一定会向低价厂商购买用于储备的商品。约束 $\lambda<1$ 意味着 $2\delta/(u-\delta)<1$ 或者

$$\delta < u/3 \tag{13}$$

约束 $\lambda>0$ 意味着 $\delta<u$，这可由式（13）明显得出。条件 $p_l+\delta=(u+3\delta)/2<u$ 与式（13）是等价的。

以正式的观点来看，到底是有 $(1-\lambda)$ 份额的厂商始终在索取低价，还是每个厂商以 $(1-\lambda)$ 的概率索取低价是没有区别的，后者是对销售行为的一个很自然的理解（反对前者的人可能会提出，当每一个人都了解价格的分布函数时，假设没有一个人确切知道哪个厂商价格较低的情况是不太合理的），当所有厂商都对定价行为采取混合策略的时候，那么在任何一个时期，索取低价格的厂商的比例就是一个随机变量，期望为 $(1-\lambda)$。然而，如果假设厂商在决定 $t+1$ 期的价格以前无法观察到 t 期价格的真实分布，那么在 t 期索取低价的厂商比例是一个随机变量的事实，就不会对分析产生实质性的影响（否则，均衡就涉及一个混合策略，（本期）低价格的大小以及索取低价的厂商比例将依赖于以前时期索取低价的厂商比例）。

3. 单一价格均衡

引理 1 已经证明了均衡中至多有两种价格，在以前的部分我们考察了两种价格情况下的均衡，现在考察一下单一价格情况下的均衡（Single Price Equilibria，SPE）我们将会证明。

定理 3：如果 $c=0$，那么 $p=u$ 就是唯一可能的单一价格均衡。并且，当且仅当 $\delta \geqslant u/3$，及 $c=0$ 时，$p=u$ 就是唯一的单一价格均衡，如果 $c>0$，就不存在单一价格均衡。

其中第一句结论可以从前面的分析中直接得出，它说明了只有当消费者为了储备购买时，$p<u$ 才可能构成单一价格均衡（SPE），但是，如果 $c=0$（并且 $\delta>0$），为了储备购买就不划算了。

第二句结论的证明是很明显的，如果 $c>0$，且 $p=u$，消费者进入市场就无利可图，如果 $c=0$，$p=u$ 要想成为均衡，当且仅当厂商为了出售储备，把价格调低到保留价格的行为是无利可图的，当所有厂商价格均为 $p=u$ 时，保留价格为 $u-\delta$。

假设这是一个均衡，那么，如果厂商降低价格的话，尽管销售量会上

升，但是利润一定会下降，因为有一半的客户是年轻人而另一半为老年人，[8] 销售会增加50%，于是我们要求有 $3/2(u-\delta) < u$ 或者 $\delta \geqslant u/3$ 成立。[9]

这个结果同定理1和定理2表明：在 $c=0$ 的条件下，如果 $\delta > u/3$，就会有 $p=u$ 这个单一价格均衡；如果 $\delta < u/3$，就会存在双价格均衡。到底是得到价格离差均衡，还是单一价格均衡，仅仅取决于存储成本的（相对）大小，如果它较低的话（相对于保留价格 u 而言），那么唯一的市场均衡一定是存在价格离差的。

如果厂商诱使消费者为了储备购买的行为是有利可图的话，那么，存储成本一定不会太高。如果存储成本不是太高的话，那么，厂商促销的行为就总是有利可图的。

为了了解为何 $c>0$ 时不存在单一价格均衡，回忆我们已经证明了当 $u>0$时，$p=u$ 是不可能构成一个单一价格均衡的。如果 $p<u$，要想构成单一价格均衡，就必然存在为了储备购买的行为。通过保留价格公式，可以很容易地证明，若是消费者为了储备购买是均衡行为的话，必然有 $c>\delta$；并且，利润最大化要求 $p=u=\delta$（消费者对是否（购买）储备无差异）。因此个人所得的消费者剩余为：$2u-2p-\delta-c=\delta-c<0$。于是没有人愿意进入市场，（这与假设）矛盾。[10]

4. 比较静态分析与福利

在双价格均衡中，价格离差的相对大小是 $u/(\delta-c)$ 的一个简单函数，令 $u/(\delta-c)=v$ 由式（10）、式（11）和式（8）可以得到：

$$\lambda = 2/(v-1) \tag{14}$$

$$p_l = (\delta-c)(v+1)/2, p_h = u \tag{15}$$

$$\bar{p} = u(v+3)/2v \tag{16}$$

---

8 回忆一下，消费者之所以要在低价的厂商那里购买储量，有两个原因：这样可以减少他的总交易成本，并且他也不确定是否能在下一期还能碰上低价。当 $c=0$ 时，只有第二种影响起作用。不过，这个影响本身，已经足以给出一个均衡的价格分布了。

9 如果 $c_1=0$（第一次的搜寻成本为零）但 $c_2>0$，就没有人愿意在第二期消费了。如果价格降到 $u-\delta$（因为只剩下年轻的消费者，在 $u-\delta$ 的价格上，他们都会购买储量），销售量就会增加一倍。所以 $p=u$ 时的单一价格均衡要求 $2(u-\delta)<u$ 或者 $\delta>u/2$。

10 如果 $c_1=0$，$c_2>0$，只要 $\delta<\min[u/2, c_2]$ 就存在一个单一价格的均衡，条件 $\delta<c_2$ 确保了不会到下一期购买。条件 $\delta<u/2$ 确保了企业不会把价格提到 u（即 $u/2<u-\delta$）。

于是有

$$\sigma_p^2 = \lambda(u - \bar{p})^2 + (1 - \lambda)(p_l - \bar{p})^2 = u^2(v - 3)/2v^2 \tag{17}$$

$$\sigma_p^2/\bar{p}^2 = 2(v - 3)/(v + 3)^2 \tag{18}$$

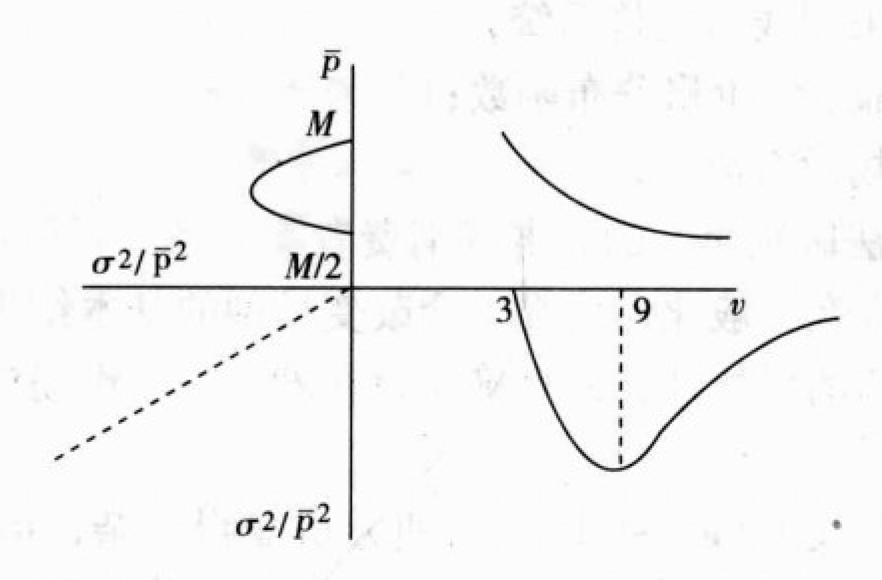


图 1　函数 $u/(\delta - c) = v$ 下的价格分布的均值和方差

图 1 表示出了均衡与 v 的值之间的关系，注意到存储成本的上升总是导致平均价格的上升——从而令促销变得缺乏吸引力——但是，它既有可能增加也有可能减少价格离差现象，增减的幅度由变异系数来衡量。

模型中的所有消费者都被设定为风险中性的，这样，他们的福利可以很容易地通过平均价格加上平均储备量以及搜寻成本来计算出来。

$$\begin{aligned}-W &= \lambda[u + \bar{p} + c] + (1 - \lambda)[2p_l + \delta] + \text{常数} \\ &= \lambda[u + \bar{p} + c] + (1 - \lambda)[p_l + \bar{p} + c] + \text{常数(利用式(1))} \\ &= 2\bar{p} + c + \text{常数(利用式(8))} \\ &= 3(\delta - c) + u + c + \text{常数(利用式(1)、式(2) 和式(10))} \\ &= 3\delta - 2c + u + \text{常数}\end{aligned}$$

因此，存储成本的增加，总是让消费者的福利受到损失，但其一般均衡效应——通过平均价格起作用——要比其局部均衡效应大得多。因此，降低存储成本（比如降低购买消费信贷的利率）就能够显著地提高福利，另一方面，降低搜寻成本的一般均衡效应会使消费者的福利降低，较低的搜寻成本会让减价促销变得丧失吸引力。

二、模型的稳健性

截至目前，我们还没有分析前面所阐述的模型所具有特殊性质：

（a）需求（效用）函数具有特殊的形式；

（b）只有两个时期；

（c）存储成本与所购买的数量成比例；

（d）所有消费者同质；

（e）厂商仅采用线性定价系统；

（f）每个人都了解价格分布函数；

（g）进入市场没有成本；

（h）厂商无法区别年轻的和老的消费者。

所有这些假设在一般化以后都不会改变模型的基本结果，确实存在着包含价格离差的均衡的可能性，并且确实存在着唯一的市场均衡也包含价格离差的可能性。

改变前 4 个假设（（a）~（d））会使分析变得复杂，但不会对结果产生根本性改变。我们在 1976 年的文章中经证明了该模型可能扩展到一个更加通常的情形，即（厂商）面临着向下倾斜的需求曲线。在 Stiglitz（1976）那篇文章中，结果被扩展到一个搜寻成本会变化的模型中，当然，把消费者存在（购买）时期扩展到两期以上是很容易想到的。这会导致均衡价格大于两种的可能性增加，在我们 1981 年那篇文章中，我们证明了是本文的结果可以扩展到其他的存储成本函数形式中去的。

剩下的四个假设（（e）~（h））需要稍作说明。

1. 搜寻成本、扰动以及均衡的存在

在以前的部分，我们假设进入成本（第一次搜寻）为零，现在我们将会证明，当它为正的时候，如果某个参数使得均衡的价格分布不存在的话，那么，它同样会使得均衡也不存在：*唯一可能的均衡必然导致价格离差。*

当搜寻是有成本的时候，消费者必须决定是否进入市场，我们会证明当不存在价格离差以及 $c>0$ 时，实际上不会有人进入市场。

此结果背后的直觉可以通过存储成本为无限大（$\delta=\infty$）这一极限情况清晰地看出来，在此条件下，每个厂商都显然是一个纯粹的垄断厂商。一旦有消费者到来，这个消费者就被完全套牢了，因为他不能再去搜寻了，于是，唯一可能的均衡就是单一价格均衡，其均衡价格为 u。然而，在那种价格下，没有消费者会愿意进入市场，因为他正确地预期到如果他进入，就会

得到负的消费者剩余。[11,12]

关于有限值的 δ 的讨论是类似的，实际上，定理的最后一部分证明了如果 $c>0$，就不存在单一价格均衡，因为（消费者）进入市场是无利可图的。[13]

因此，一旦不存在有价格离差的均衡，也就不存在均衡了，“*残酷的竞争加上一点点垄断力量就毁掉了整个市场均衡*”。

存在着这样一种价格分布，它是市场内在扰动的一种形式。通过它可以部分地解决这个问题。[14] 扰动的存在保留了消费者遇到一桩好买卖的希望，正是这种希望使消费者愿意进入市场。

我们通过扩展定理 2 可以建立：

定理 4：均衡存在的充分和必要条件是 $u \geqslant 3\delta - c$ 以及 $\delta > c$。当均衡存在时，则由定理 1 中的描述可以刻画出均衡。

由式（11），这要求满足 $0<\lambda<1$ 和 $\delta>c$。如果 $u>3(\delta-c)$，那么这些条件显然都得到满足。不走运的年轻人在年老后再次进入市场，有 $\bar{p}+c=p_l+\delta \leqslant u$，但由式（10）可以得到 $(u+\delta-c)/2+\delta \leqslant u$ 或者 $u \geqslant 3\delta - c$。如果不走运的年轻人在年老后再次进入市场是有利可图的话，那么，他在第一期就进入市场也一定是有利可图的。

2. 如果垄断者实行价格歧视，均衡就不会存在

关于厂商无法区分一个年轻的消费者和一个年老的消费者的假设，对于前面的分析是非常关键的。要是厂商可以完全地实施价格歧视，那么，它们

---

11 正式地说，尽管为了诱使消费者进入市场，需要把价格定在 $u-c$，但是对于每个小企业——只要它确信对于市场平均价格的影响可以忽略不计——就必然有提价至 u 的激励；因为一旦有消费者（不走运）来惠顾，这个消费者就被套住了，并且会支付价格 u（过去的事就让它们过去吧：虽然这个消费者也很后悔来这个店，但是他不能陷于后悔，而必须在不利情况下尽力而为）。

12 假设 $\delta=\infty$ 就消除了未来与现在的任何联系，从而可以从静态来考察市场。彼得·戴梦德（Peter Diamond，1971）模型可以视为我们的模型在 $\delta=\infty$ 条件下的特例。但是，戴梦德没有检验实际中消费者进入市场究竟是不是有利可图的。而我们的分析（基于特定的效用函数和搜寻成本函数形式）证明那绝对是无利可图的。并且，我们还证明了如果企业可以实行非线性定价策略，那么不管消费者的效用函数是什么形式，只要存在正的搜寻成本，就不会有均衡存在。

13 如果第一期的搜寻成本 c_1 小于第二期的成本 c_2，那么只要 $\min[u/2, c_2] > \delta > c_1$ 并且厂商不能够采取价格歧视，就一定会有 $p=u-\delta$ 的单一价格均衡（这个结果将脚注 10 的结果一般化了）。

14 这个结果同（Grossman and Stiglitz，1980）的研究结果是相似的。在他们对资本市场的分析中，由于存在信息成本，扰动对于确保均衡的存在是至关重要的。但是，与本文的模型相反，那里的扰动完全是外生的。

就会把价格提到足够高，榨取完所有的消费者剩余，只要有消费者来的话（对于老年消费者，索取价格为 $p=u$；对于年轻消费者，由于他要买两个单位，因此，总的价格为 $2u-\delta$）。如果 $c_1>0$（c_1 为第一次进入市场的交易成本），那么，消费者进入市场都是无利可图的，也就不存在均衡了。[15]

要是厂商可以实施非线性定价策略，那么它们总是能够确认出谁是年轻人，[16]因为只有年轻人愿意购买两个单位。厂商要做的不过是给一个足够大的数量折扣，使得相对于再次进入市场而言，年轻的消费者更愿意（购买）储备。[17]这样，我们能够证明如果搜寻成本存在的话，均衡是不存在的。

3. 理性预期

我们假设了消费者确实知道价格的分布情况，尽管他们不知道任何一家厂商的价格。这是一种“理性预期”式的假设。之所以用它，倒不是因为它的现实性，而是由于要找一个替代的关于预期的假设十分困难。

然而，很明显，这个假设预期是十分关键的。因为消费者决定是为了储备购买，还是只购买当期消费量，都取决于他对于再次进入市场后要购买物品价格的预期。认为消费者关于价格的概率分布预期是与价格的真实分布相吻合的观点，缺乏让人信服的理由。实际上相当数量的文献都曾经提出，个人的预期概率分布可能存在着系统性的误差，尤其是当事件（比如打折促销）出现并不频繁的时候（Amos Tversky and D. Kahneman，1974）。

厂商愿意把价格定在低于垄断价格水平的地方是因为，如果它们定价足

15 为了得到这个结果，首先要证明消费者绝对不会在第二期再进入市场。假设消费者真的还会在第二期进入市场，那么他在第一期其实只需要购买一个单位。假设 $\hat{p}$ 是购买第二单位的保留价格。很明显，只要 $\hat{p}>0$，那么以 $\hat{p}$ 的价格再出售一单位的商品就是有利可图的。因此，第一期的企业就不是利润最大化的了。于是，唯一的均衡必然使得消费者在第一期就购买两个单位的商品。但是，如果所有的消费者都在第一期就购买两个单位的商品，均衡必然要求企业对这两个单位的商品的要价为 $2u-\delta$，而此时的消费者剩余又会是负的（这由以下的事实可以迅速地推出，即企业对老年人的要价都必然等于 u）。

16 Salop（1977），Stiglitz（1977），M. Katz(1981）探讨了有关采取非线性定价策略的问题，其中垄断企业可以对不同的消费者实施部分的价格歧视。

17 这种定价方法是很容易计算的。令 $R(Q)$ 为消费者为了得到 Q 数量的商品的支出。显然有 $R(2)\geqslant R(1)$ 并且有 $R(1)\geqslant u$（假设 $R(1)<u$，那么提高 $R(1)$ 到 u 可能仅仅只是损失了本来要购买两个单位年轻的客户。但是，由于 $R(2)\geqslant R(1)$，就不会有损失）。注意到无论其他企业的价格是如何制定的，$R(2)\geqslant R(1)\geqslant u$ 都必须成立；所有企业的价格方案都必须满足这些性质。但是，$R(1)\geqslant u$ 在消费者存在的两期均成立，如果 $R(2)=2u-\delta$，消费者就会在第一期购买两个单位的商品。这就是利润最大化的值 $R(2)$。这个收入方程使得消费者剩余为零，因此，如果进入市场是有成本的话，就没有人进入了。

够低，消费者就会为将来而消费。在一个存在价格离差的经济中，存在双价格均衡要求下式成立：

$$\hat{p}(2+\lambda) = u(1+\lambda) \tag{19}$$

消费者的保留价格依赖于他们对于未来商品平均价格的预期：

$$\hat{p} = \bar{p}^{e} + c - \delta \tag{20}$$

而$\bar{p}$的上标e，提醒我们关键是消费者对于平均价格的预期。正是这些预期价格影响了打折促销的利润率，接下来又影响了打折促销发生的频率；但同时，厂商（促销的利润率和频率）也反过来影响了平均价格，平均价格又会反过来影响预期价格。

假设消费者相信存在着折扣促销，即使（折扣促销）根本没有发生，唯一的均衡也一定包含一个价格分布。比如可以假设消费者的预期价格高于其真实的平均值，$\bar{p}^{e} = \beta(\bar{p})\bar{p}$，$\beta > 1$，并且对$\bar{p} = u$有$2\bar{p}/3 < \beta(\bar{p})\bar{p} + (c-\delta)$成立。即使在理性预期这样严格的假设条件下也能够存在的单一价格均衡，在此条件下竟然不存在了。因为倘若假设均衡存在，同从前一样，假设$c=0$且$\delta > u/3$，会有一个$p=u$的单一价格均衡存在。如果$p=u$，那么任何把价格降到$\hat{p} = \beta(u)u + (c-\delta)$的厂商都会增加50%的销售量。因此，如果在$\bar{p}=u$，有$\{\beta(\bar{p})\bar{p} + (c-\delta)\}3/2 > u$，那么降价促销就是有利可图的（显然，如果$\beta > 1$，不等式$u/3 < \delta < (\beta - 2/3)u$会同时得到满足）。

因此，只要消费者对于折扣促销的预期不是过分地悲观，那么均衡时，就一定存在折扣促销。

三、结论

就特定的市场而言，在传统经济分析中，传说中的拍卖人也许是描述市场如何运行的一个很好的原则。然而对于其他的市场而言，要是真是存在拍卖人的话，其所达到的均衡就会与真实均衡大为不同。本文正是描述了在可能是最简单的条件下，即存在信息采集成本条件下，两者会有怎样的不同。

我们模型的基本结构是很简单的。企业决定自己的价格。由于存在信息采集的成本，企业面临着向下倾斜的需求曲线，而需求曲线又反过来取决于其他企业的定价。我们可以把第i个企业的利润写为价格向量的函数：

$$\pi_i = \pi_i(p_1, \cdots, p_N)$$

纳什—古诺均衡可以由下列两组条件来刻画：（a）市场现存的所有企

业都制定利润最大化的价格；（b）对于所有市场现存的企业，利润都是非负的，并且对于任何潜在企业，在所有潜在价格条件下，利润是非负的。

如果所有企业都是同质的，并且企业数量足够大的话，假设均衡中存在价格离差，那么利润函数（作为价格的函数）就一定在每一个价格处达到最大值，并且每个最大值都相等。[18]一般有两大类机制可以令价格不同的企业得到相同的利润。第一类机制是，价格低的企业销售量大，并且正好补偿了低价的损失。这种现象可以有很多种原因。销售量的上升可能是由于更多的顾客光顾，可能是由于有更多顾客愿意购买，或者由于消费者愿意购买的数量更多了。第三个可能性在本文中有详细的探讨。第二类机制是，如果企业必须招揽才会有顾客的话，比如，做广告，那么价格高的企业可能其招揽单位消费者的成本也高，正好抵消了较高的单位利润。[19]

我们在分析中一共确认了三种均衡情形。其中每一种情形都与传统所分析的完全竞争市场均衡大不相同（尽管都允许市场的自由进入）。

（a）均衡可能是存在价格离差的，这里个人所支付的价格高低仅仅依赖于他的运气。模型不仅可以直接应用于某一期出现的价格离差现象（同时存在着高价厂商和低价厂商），也可以直接应用于分析折扣促销现象（未做广告的特价品）。我们注意到价格的变化同样与质量的变化相关，比如耐用性；因此，重要的是（分析）单位有效商品的价格。

（b）均衡也可能只包含单一价格，但这个价格会高于完全竞争市场均衡的价格。

（c）均衡也可能不存在。当允许厂商实行数量折扣时，均衡就总是不存在的。

竞争通常被认为是对消费者有利的事情；尽管我们都不反对这个一般假定，但是，对于竞争的另一面我们也必须牢记于心：当存在搜寻成本时，竞争可能采取这样的形式，企业总是试图更好地利用由于搜寻成本和信息所赋予的小小垄断力量（更成功的企业可能不是因为它们更有效率，而是因为它们更能够有效实施价格歧视）。虽然在没有交易成本的条件下，我们都知道实行完全价格歧视的垄断厂商是帕累托最优的，但是，一旦引入交易成本，完全竞争市场均衡就不复存在了，（此时）没有人愿意进入市场。

---

18 如果企业数目很少，均衡的性质就依赖于潜在的进入者是如何看待他进入的后果的（Salop，1977）。同样地，厂商也必须推测价格如果改变的后果（参见我们在 1977 年的文章）。

19 Gerard Butters（1977）。类似的论证方法可以应用于质量差异的分析。高质量的企业回头客就多，这样他们的平均招揽顾客的成本就低。

本文的模型还具有另外一层道德含义：由于信息成本的存在，以往那些将传说中的市场经济与传说中的社会主义经济相比较而得到的比较制度分析结果的重要性可能十分有限。迄今为止的成果，比如几条基本等价性定理，很有可能会产生严重的误导。在我们所建立的模型中，市场产生了扰动。并且由于信息成本的存在，这种扰动是无法消除的。尽管我们还没有构造出与完全竞争经济或者计划经济非常吻合的模型，但至少有一点是很明显的，就是这两者可能会有显著的区别。

参考文献

Butters, Gerard, "Equilibrium Distributions of Sales and Advertising Prices," *Review of Economic Studies*, October 1977, 44, 465 – 91.

Diamond, Peter, "A Model of Price Adjustment," *Journal of Economic Theory*, June 1971, 3, 156 – 68.

Grossman, Sanford and Stiglitz, J. E., "On the Impossibility of Informationally Efficient Markets," *American Economic Review*, June 1980, 70, 393 – 408.

Katz, M., "Imperfect Competition and Heterogenous Consumers: The Theory of Screening in Product Markets," unpublished doctoral dissertation, Oxford University, 1981.

Perloff, Jeffrey and Salop, Steven, "Equilibrium with Imperfect Information and Product Differentiation," Federal Trade Commission, 1980.

Salop, Steven, "Monopolistic Competition with Outside Goods," *Bell Journal of Economics*, Spring 1979, 10, 141 – 56.

________, "The Noisy Monopolist: Imperfect Information, Price Dispersion, and Price Discrimination," *Review of Ecomomic Studies*, October 1977, 44, 393 – 406.

________ and Stiglitz, J. E., (1976a) "Search Costs, Monopoly Power, and Price Distributions," mimeo., Stanford University, 1976.

________ and ________, (1976b) "The Theory of Sales, or Luck of the Draw: A Simple Model of Equilibrium Price Dispersion with Identical Agents," mimeo., 1976.

________ and ________, "Bargains and Ripoffs: A Model of Monopolistically Competitive Price Dispersion," *Review of Economic Studies*, October 1977, 44, 493 – 510.

________ and ________, "The Theory of Sales: A Simple Model of Equilibrium Price Dispersion with Identical Agents," Econometric Research Program, research memo. 283, Princeton University, June 1981.

Steiner, Robert, "Understanding the Consumer Good Economy," mimeo., 1978.

Stiglitz, J. E., "Monopoly, Non-linear Pricing, and Imperfect Information: The Insurance Market," *Review of Economic Studies*, October 1977, 44, 407 – 30.

________, "On Search and Equilibrium Price Distributions," in Michael Boskin, ed., *Economics and Human Welfare: Essays in Honor of Tibor Scitovsky*, New York: Academic Press, 1979, 203 – 16.

Tversky, Amos and D. Kahneman, "Judgment under Uncertainty: Heuristics and Biases," *Science*, 1974, 185, 1124 – 31.

市场竞争与企业数量：双头垄断比小摊市场更具有竞争性？*

本文运用标准搜寻模型的变形分析了市场均衡以及增加市场之中企业的数目对于市场均衡的影响。如果边际搜寻成本随着搜寻次数的增加而增加的话，那么企业的需求曲线一定是弯折的，并且价格下降时的需求弹性要小于价格上升时的需求弹性；当边际成本变化时价格可能不发生相应的变化。随着市场中企业数目的增加，均衡条件下的最高价格水平会逐渐增加，接近垄断价格，但是其最低价格水平却降低了。另外，如果边际搜寻成本随着搜寻次数的增加而降低的话，那么均衡如果存在的话，其中就存在着价格差异。

到底需要多少个企业才能够使市场的竞争功能完全发挥出来？如果存在着完美信息，那么即使市场之中只有两个企业，它们也会像伯川德（Bertrand）竞争者那样将价格定在边际成本处。[1] 相反，在古诺均衡条件下，随着企业数目的增加，需求弹性会上升，而价格会下降，只有在极限情况下，才可能达到完全竞争价格。

虽然在很早以前，人们就认识到不完美信息会造成市场运行的竞争性下降（Scitovsky，1950；Salop，1976），现代经济学理论甚至开始认真审视这一结论——增加市场企业数目会导致价格下降——在定性结果方面的稳健性。尽管目前对有限企业数目的市场成本均衡性质方面的研究还不是非常细致，但是对于无限企业数目的市场均衡性质的研究却是非常细致的。如果工人能够续贯地搜寻而且搜寻成本是有限的（无论成本多么的小），那么市场均衡价格就一定是垄断价格。我们还得到了另外两个令人吃惊的结果：（1）

---

* "Competition and the Number of Firms in a Market: Are Duopolies More Competitive than Atomistic Markets?"，本文受斯坦福大学社会科学数量研究所国家科学基金的资助，项目号 SES－8320464，并且还受到胡佛研究所和普林斯顿大学国家科学基金项目的资助。多年以来，同 Steve Salop 和 Michael Rothschild 关于本文题目的讨论，令我受益匪浅，而 Peter Crampton，José Scheinkman，Bernard Salanie 以及 Arunava Sen 对于本文初稿的评述是非常具有启发意义的。atomistic market，小摊市场，指有很多小企业的市场，又译为"蚂蚁市场"、"摊贩市场"或"原子市场"——译者。

1 假设这里不存在产能约束。

不存在均衡的价格差异（Diamond，1971）；（2）如果第一次搜寻是有固定成本的，那么当（a）企业能够采取非线性定价策略；或者（b）个人的效用函数导致即使价格水平低于某个保留价格水平，他们也只够买一个单位的时候，均衡是不存在的。

搜寻的成本，就算很小，也会对均衡的性质产生巨大的影响。[2]

本文的目标是考察三种对于标准的续贯搜寻模型进行微小修改的结果：（1）存在着某些个人的搜寻成本是无穷小的情况，如果由于搜寻成本 s 产生的分布函数 $F(s)$，$f(s)$ 是其密度函数，那么 $f(0)>0$[3]；（2）企业数目是有限的而不是无限的；（3）增加搜寻一个企业的成本可能依赖于已经搜寻过的企业数目。

如果市场企业数目是有限的，均衡价格就会低于垄断价格：双头垄断看起来是要比市场企业数目无限多的情形更加具有竞争性。这是由于在已经研究过的无限市场企业数目模型之中，如果某个企业降低价格，也不会引致搜寻行为，而在双头垄断的情况下，则一定会引起搜寻行为。因此市场价格弹性在双头垄断的情况下，要大于无数企业的情况。随着企业数目的增加，找到一个低价企业的成本是上升的。因此引致搜寻（induced search）数量是下降的。增加企业数目的效果，看起来是非竞争性的。

但是，随着企业数目的增加，会产生两种效应：如果企业降价，只有那些搜寻成本相对较小的个人才愿意搜寻，因为在此情况下搜寻到这个“减价机会”的难度是增加了的。尽管任一企业的一小部分消费者出现引致搜寻，但是其他企业那里存在着更多的值得吸引的消费者。当搜寻成本是常数，并且存在着重置搜寻（search with replacement）的情况下，[4] 可以证明第二种效应要超过第一种效应：传统的认为市场的竞争程度会随着市场之中的企业数目增加而上升的结论就需要修正了。在极限的情况下，随着市场之中的企业数目增加到无穷大，市场价格从垄断价格水平单调下降直到某个水平，降低的程度依赖于 $f(0)$，前提只要 $f(0)>0$。[5]

2　想查询有关结果，参见 Stiglitz（1979）。还有 Axell（1977）和 von zur Muehlen（1980）。

3　Braverman（1980）；Braverman and Dixit（1981）考察了 $f(0)>0$ 情况下的结果，这与 Salop and Stiglitz（1977）的搜寻模型的条件略有不同。

4　不存在重置搜寻是指，如果某个人搜寻了某个企业并且拒绝了该企业，那么他是不会再回到该企业的。

5　在极限的情况下 $f(0)=0$，存在着某个企业数目 N，当低于它的情况时是不存在单一价格均衡的（Arrow and Rothschild，1973）。

因此在$f(0)>0$的条件下，以市场之中的企业数目来反映价格的垄断水平的标准均衡模型是不正确的。

标准模型的解释尽管考虑到了第一种效应，但是没能考虑到第二种效应。因此，随着市场之中的企业数目增加，企业从数目众多的别的某个企业那里能够吸引的消费者数目是很少的。当然，就认为企业从别的某个企业那里能够吸引的消费者数目趋近于零的速度，是要比市场之中的企业数目增加速度更快，也是不正确的。

标准模型的解释的另外一个也是非常重要的错误在于：它没有考虑在企业找到消费者的时候，市场所产生的信息不对称。消费者仅仅知道该企业的价格，以及其他企业的价格分布。这种信息不对称会产生严重的后果：价格上升的需求弹性可能会与价格下降的需求弹性有很大的区别。一般情况下，需求曲线是弯曲的。这样又会产生两种结果：(1) 如果企业的数目是有限的，边际搜寻成本是常数，也不存在重置搜寻，那么唯一的市场均衡会产生价格差异。在边际搜寻成本是递减的情况下，这个结论也是成立的。此结果是对萨勒普和斯蒂格利茨（Salop and Stiglitz，1977）的分析结果的概括，在分析中他们仅仅考虑了搜寻技术是凸的有限的情况，并且在第一次搜寻之后的所有搜寻事实上都是无成本的。(2) 如果边际搜寻成本是递增的，那么均衡可能就不能确定了，而产量可能会随着边际搜寻成本的变化而变化。

之所以要考察能够导致价格差异的模型，是有两个原因：首先，大量的证据表明，市场中同质商品的出售价格常常是有差异的，[6] 这在产品市场上至少是可能出现的。[7] 其次，我在以前曾经提到过，在可行的条件下，如果第一次搜寻是有成本的，那么不存在市场均衡。如果商品售价低于垄断价格的可能性至少是存在的话，那么这个难题是能够解决的。因此我们所建立的模型是内在一致的，而这是假设所有消费者都具有严格为正的搜寻成本，以

---

6 那些相信单一价格法则的人可能会声称，不同商品售价不同是源于其他方面的差异，比如，地点、服务等。在某些情况下，这可能是对的。但是在其他的情况下，商品的售价差异是如此之大，以至于仅仅靠质量差异是无法解释全部的价格差异现象的。在质量差异为主导的市场之中，最适合的市场分析模型就是区分商品的市场搜寻模型。这些市场一般情况下是很自然地会产生价格差异的，就算商品的生产技术是同质的，它们的售价也会有差异。

7 实际上，这正是 Rothschild（1973）对于 Stigler（1961）的模型最为突出的批评：给定斯蒂格勒（Stigler）关于消费者行为的假设，所有企业都会索取相同的价格。

及存在着企业连续统的产品市场模型所不具备的。[8]

本文一共分为四个部分。第一部分给出了模型，而第二部分给出了有限企业数目和在线性搜寻成本技术条件下的结果。第三部分证明了在非线性搜寻成本技术条件下分析会有什么不同。第四部分分析了结果的稳健性。

一、模型

为简单起见，考虑一个市场，其中对于给定消费品 $x = x(p)$，所有人都具有相同的需求函数，但是他们的搜寻成本是不同的。个人的间接效用函数可以写作

$$U = v(p) + Y \tag{1}$$

其中 Y 是个人的收入，因此由罗尔公式

$$x = -v' \tag{2}$$

个人的搜寻规则是很简单的：如果在第 t 次搜寻情况下有 $p \leqslant \hat{p}_t$，就购买；反之就不购买。$\hat{p}_t$ 项是第 t 种商品的保留价格水平。市场价格水平正好使得消费者从某企业购买商品的期望效用，等于继续搜寻所得到的效用水平。继续搜寻所得到的效用水平的高低依赖于搜寻过程，例如，过程是否存在着重置搜寻 。[9]

假设市场上存在着 L 个消费者以及 N 个企业，[10]每个消费者都知道价格的概率分布，但不知道各个企业的价格。[11]消费者在企业之间进行随机的搜寻。每一个消费者都会不断地搜寻，直到找到一个售价等于或者低于他的保留价格水平的企业。

记 s_t 为第 t 次搜寻的成本，以往的文献主要关注于三个特例：（a）当 $t>1$时，$s_t = \bar{s}$ 为常数的线性搜寻成本函数情况。由于在 $s_1 > 0$ 条件下均衡可能不存在，因此我假设 $s_1 = 0$。（b）凸性搜寻成本函数的情况。时间（以及

8　正式地说，我们可以通过假设第一次搜寻的成本为零这种标准方法来回避均衡不存在的情况。如果我们假设第一次搜寻的成本和以后的搜寻的成本相同，那么只要平均均衡价格足够小，即 $f(0)$ 足够大，我们的分析结果就不会受到根本影响（否则，我们又会碰到均衡存在性问题）。

9　假设搜寻成本是常数，如果消费者今天搜寻了一家售价为 p 的企业之后，还继续搜寻是有利可图的话，那么在这两种情况下，消费者另外搜寻到一家售价为 p 的企业之后还继续搜寻也是有利可图的。

10　整个分析过程之中，企业的数目都是外生决定的，但是在企业的平均成本是 U 形的条件下，实现 N 的内生化是十分容易的，见 Salop and Stiglitz（1977）。

11　这是文献之中的标准假设。尽管某些情况下它可能是正确的，但我们主要是为了分析简单才如此假设的。

其他资源）的稀缺性意味着搜寻成本可能会随着搜寻次数的增加而增加：对于所有的 t 有 $s_t \geqslant s_{t-1}$。它的一个特例是（Salop and Stiglitz，1982）曾经使用过）：在某个 t 值之前搜寻成本是固定的，一旦 t 超过这个值，搜寻成本是无限大的。（c）凹性搜寻成本函数的情况。如果企业是分散分布于某个商业区之内的，那么某次购物的固定成本主要就在于抵达该商业中心的成本。同样地，如果消费者能够通过诸如消费者报告等来获取信息，那么他们就可能知道当时一些企业的价格。这些都是边际搜寻成本小于平均搜寻成本的情况，即对于所有 t 有 $s_t \leqslant s_{t-1}$。这个例子的极限情况（Salop and Stiglitz（1977）分析过该情况）是当 $t>1$ 时，有 $s_1>0$，$s_t=0$。

各个消费者的搜寻成本是不同的。在线性搜寻成本函数情况下，我们可以用分布函数 $F(s)$ 来表示它。

企业的收入可以表示为

$$R = Mpx \tag{3}$$

其中 M 是购买了商品的消费者的数目，M 是某企业的售价和其他企业的售价的函数。产品的边际成本是 c，于是利润最大化企业将使得产品的边际成本等于其边际收益。

$$p = \frac{c}{1-(1/\epsilon)} \tag{4}$$

其中 ϵ 是企业需求曲线的弹性

$$\epsilon \equiv -\frac{d\ln M}{d\ln p} + E(p^*) \tag{5}$$

而 E 是个人需求曲线的弹性

$$E(p) \equiv -x'(p)\frac{p}{x} \tag{6}$$

注意到对于一个垄断厂商，有 $d\ln M/d\ln p=0$，因此 $\epsilon = E(p^*)$ 我们把这个价格水平记为垄断价格水平。

$$p^m = \frac{c}{1-[1/E(p^m)]}$$

当 M（即消费者的数目）不是 p 的可微函数的时候，我们分别用 ϵ^+ 和 ϵ^- 来标记价格上升和价格下降时的需求弹性。

如果 $\epsilon^+(p^*) > \epsilon^-(p^*)$，如图 1 所示，那么需求曲线在 $p=p^*$ 处发生弯折，但是只要满足

$$\frac{c}{1-[1/\epsilon^+(p^*)]} < p^* < \frac{c}{1-[1/\epsilon^-(p^*)]} \tag{7}$$

价格最后是不会随着边际成本的变化而变化的。

本文的剩余部分大多是用于讨论消费者数目是如何随着价格 p 的变化而变化的；即，M 是否是可微的以及 ϵ^+ 和 ϵ^- 的值究竟是多少？

二、有限企业数目，$f(0)>0$，以及在线性搜寻成本函数情况下的对称均衡

当市场之中的企业数目是有限的，是否存在着重置搜寻就会对均衡产生影响。[12]在本部分的前两点，我们只考虑对称均衡的情况。在线性搜寻成本函数情况下，保留价格水平 $\hat{p}_t$ 是个常数，记为 $\hat{p}$。

1. *不存在重置搜寻*

考虑一个市场之中有 N 个企业，售价均为 $p=p^*$，在均衡条件下，每个企业的消费者数目为 L/N。如果有个企业考虑提高售价 $p>p^*$，那么它的消费者之中那些搜寻成本较低的将会离开，因为消费者们都确信，如果再次搜寻一定可以找到价格更低的企业。于是，所有搜寻成本低于 $\hat{s}$ 的消费者都会选择离开。而 $\hat{s}$ 由下式定义：

$$v(p^*) - v(p) = \hat{s}(p) \tag{8}$$

式（8）告诉我们，究竟会有多少人会进行搜寻：所有 $s<\hat{s}$ 的消费者。函数 $F(\hat{s}(p))$ 给出了在每一个 p 值条件下搜寻人数所占比例。利用式（2）可以得到：

$$\frac{d\hat{s}}{dp} = -v'(p) = x(p) \tag{9}$$

企业提价之后发现消费者数量仅仅为

$$M = \frac{L}{N}[1 - F(\hat{s}(p))] \tag{10}$$

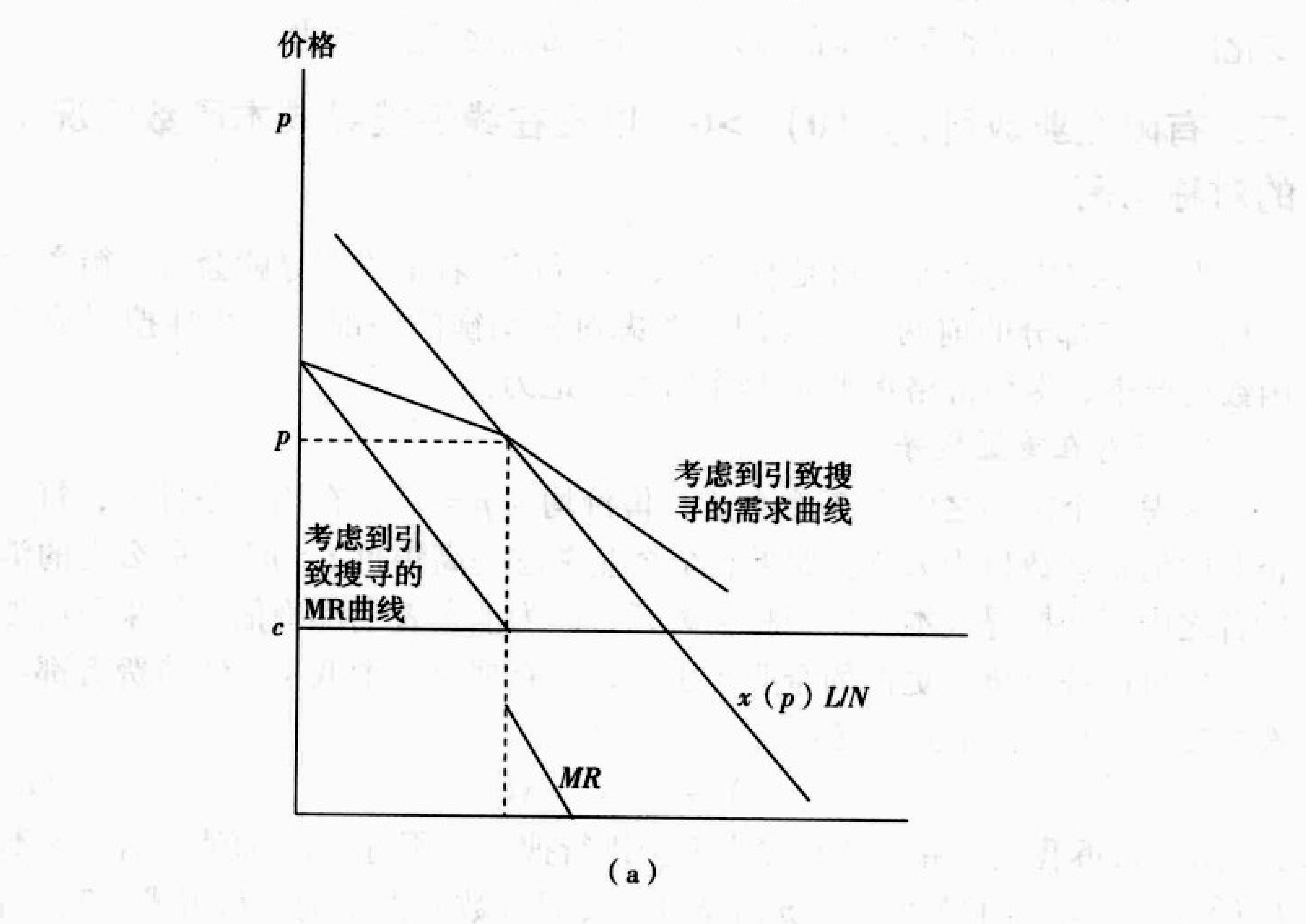
将式（10）微分可以得到 $p=p^*$ 处的需求弹性：

$$\epsilon^+(p^*) = E(p^*) + f(0)x(p^*)p^* \tag{11}$$

考虑将价格降低到 p^* 以下的企业了解到，搜寻成本较低的消费者将会从其他企业那里被吸引过来，而且搜寻成本较高的产生的消费者也不会离去。这样他的期望消费者数目就是

$$L\left[F(\hat{s}(p)) + \frac{1 - F(\hat{s}(p))}{N}\right] \tag{12}$$

12　此假设的可行性依赖于搜寻者的记忆以及企业是否采取了混合策略。后面我们会讨论这一点。



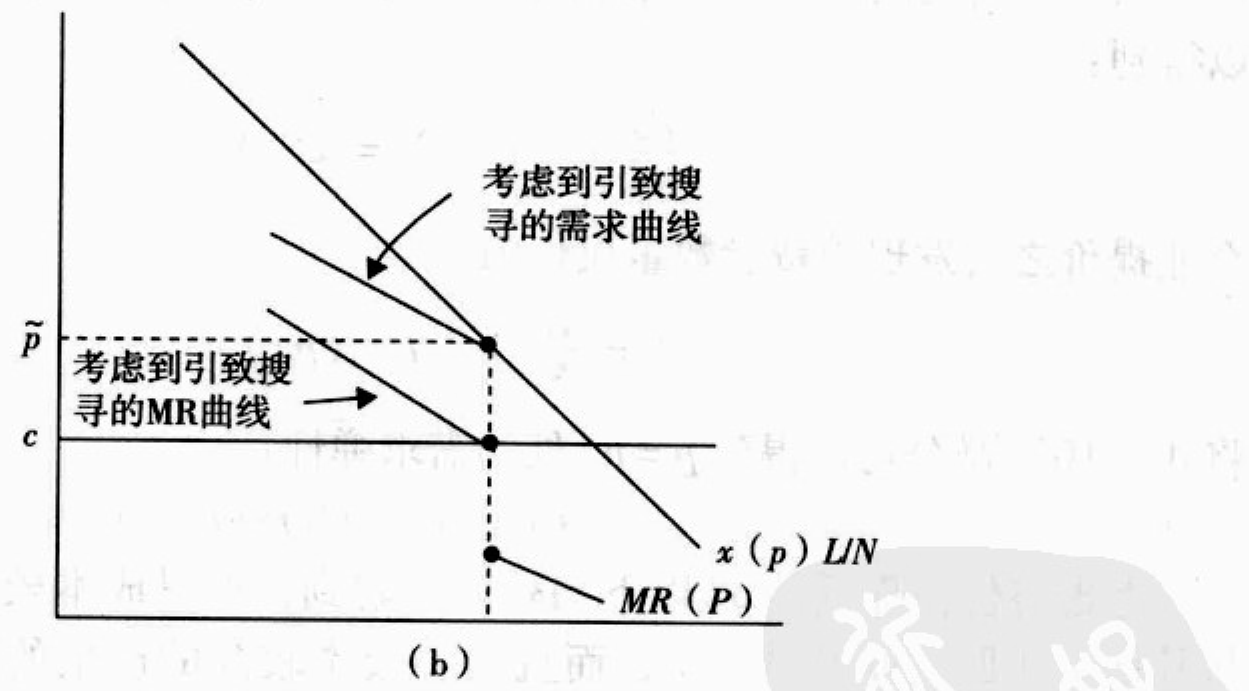


图1（a）弯曲的需求曲线

（b）在企业连续统和 f（0）>0 条件下的均衡：

均衡是不确定的，价格必须低于或等于垄断价格，但高于或等于 $\tilde{p}$

其中 $\hat{s}$（p）就是使得所有 $s \leqslant \hat{s}$ 的消费者不断地搜寻直到找到售价较低的企业为止的搜寻成本（它是售价较低的企业售价的函数）。

希望找到一个低价企业的期望搜寻者数目为 $N/2$，因此 $\hat{s}$ 由下式给出

$$v(p) - v(p^*) = \hat{s}(p)\frac{N}{2} \tag{13}$$

可以直接得到在价格上升条件下的需求弹性为

$$\epsilon^-(p^*) = E(p^*) + \frac{2(N-1)}{N}f(0)p^*x(p^*) \geqslant \epsilon^+(p^*) \tag{14}$$

等于号仅仅在 $N=2$ 时取得。因此如果存在着两个以上的企业，需求曲线就会发生弯折，在价格降低条件下的需求弹性就会超过价格上升条件下的需求弹性。这就意味着对于 $N>2$ 是不存在单一价格均衡的（见图2）。

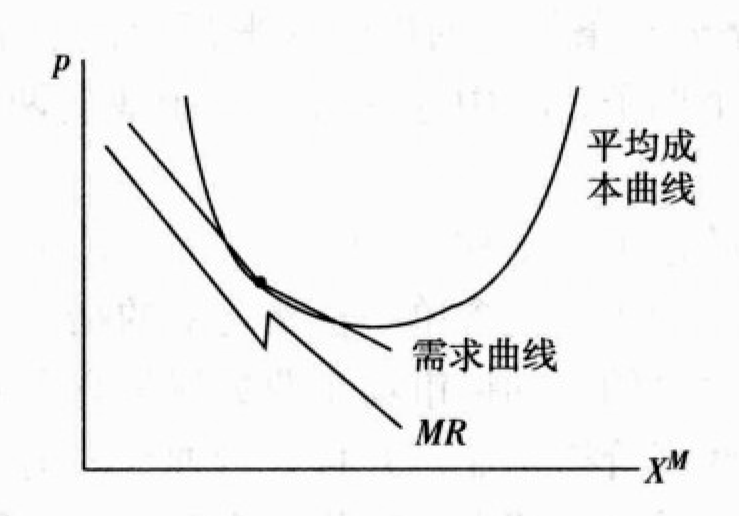


图2 在三个企业以及无重置搜寻条件下，在 p^* 点有弯折，因而不会存在一个单一的价格均衡

2. 重置搜寻

这里的分析过程和以前非常相似，只不过如果某个企业想要提高价格的话，会使得随机抵达这个高价企业的消费者平均花费 $N/(N-1)$ 次搜寻来找到价格更低的企业。然而，如果某个企业想要降低价格的话，会使得随机抵达某个高价企业（售价为 p^*）的消费者平均花费 N 次搜寻来找到这个价格更低的企业。只要修改一下式（13）使得边际消费者以适当的方式来进行搜寻，我们就可以迅速得出价格上升以及价格下降时的需求弹性：

$$\epsilon^+ = \epsilon^- = \epsilon = E + f(0)px\frac{N-1}{N} \tag{14'}$$

这里需求曲线是不存在弯折的；并且存在着单一价格水平的均衡。另外，随

着企业数目的增加，市场价格会单调下降。[13]

在极限条件下，市场价格低于垄断价格水平 p^m

$$p = \tilde{p} = \frac{c}{1-(1/\tilde{\epsilon})} < \frac{c}{1-[1/E(p^m)]} = p^m$$

其中

$$\tilde{\epsilon} = E(\tilde{p}) + f(0)\tilde{p}x(\tilde{p}) \tag{15}$$

3. 均衡价格分布

当不存在重置搜寻的情况下，是不存在单一价格均衡的；在唯一的均衡里不同企业的售价是不同的。而即使在存在单一价格均衡的条件下，还是可能存在着差异市场价格的均衡的。[14]由于每一个企业都追求将需求弹性置于超过 E 的水平（如果它提高价格，需求会降低；如果降低价格，需求会上升），在所有均衡价格分布条件下的价格水平都会低于垄断价格水平。在附录Ⅰ里我们给出了一个例子，其中存在着三个企业有两个收取较高的价格，一个收取较低的价格。[15]

在附录Ⅰ里给出的这个例子和 $f(0)=0$ 的条件是一致的。因此在有限搜寻成本条件下唯一的均衡，包含单一价格均衡的结论，是极其依赖于市场企业数目无限的这个假设的。如果市场企业数目是有限的，那么当低价的企业提高售价时，对那些高价格企业，并且进行搜寻的消费者会发现搜寻是不值得的。消费者数目不受价格的微小变化的影响这个性质，对于唯一的均衡

13 此结果同 Satterthwaite（1979）的研究是有着非常大的差异的，那里他提出随着企业数目的增加，不完美信息的程度会加剧。这样会降低需求弹性，从而增加市场均衡价格（Pauly and Satterthwaite，1981）。他的模型同这里的模型有两个重要区别：（a）他考虑了异质商品的情况，异质商品不同特性对于不同的消费者而言有着不同的价值；而我们的模型之中商品是同质的；（b）他所假设的信息获取技术也与这里的假设有着很大的不同。他假设个人是通过一种信誉机制来获取信息的而非搜寻。尽管他的模型的假设与有关消费者所拥有的信息以及他们怎样获取更多的信息的某些替代性假设是一致的，但是却与这里所给出的续贯搜寻模型的假设不一致。Salop（1979）以及 Salop and Stiglitz（1987）也注意到信息结构的改善，可能增加也可能减小需求弹性。因此不同的假设条件导致结果产生如此大的差异也就不足为奇了。Rosenthal（1980，1982）给出了一个完全不同的模型，其中随着企业数目的上升售价会上升。

14 假设所有企业的生产技术相同，并不意味着它们就会索取相同的价格（Butters 1977；Salop and Stiglitz，1977，1982；Satterthwaite，1979；Stiglitz，1985）。相比而言，在技术条件不同的情况下建立差异价格模型要更加容易（Reinganum，1979）。

15 通常，有两种对于价格差异的理解：一种含了混合策略，而另一种假设某些企业永远索取高价而其他的企业永远索取低价。前者的优点在于假设消费者知道价格的整体分布，而不知道个别企业的价格水平是更加可信的。在此条件下恰当的搜寻模型应该包含重置搜寻，只要搜寻的时间和企业保持售价的时间能够吻合。

里存在单一价格是十分重要的。

4. 企业的连续统

如果存在着企业的连续统（continuum of firm），所有的售价一定会产生相等的利润；每一个企业的规模都足够小，以至于它们认为不会对市场的搜寻行为产生影响，因此在某个价格水平上会达到利润最大值，所有企业都会把价格调整到利润最大化的价格水平。任何价格分布都会产生特定的搜寻行为（个人的保留价格随着搜寻成本的变化而变化），这就能够决定在每一个价格水平上的销售量以及相应的利润水平。问题在于，是否存在着一个非退化的价格分布，使得对于每一个价格水平和利润水平都相等？在附录Ⅱ里，我证明了这是可能的；我推导出了价格分布必须满足的微分方程，并且给出了一个例子。我还证明了分布之中的最高价格水平要低于垄断价格水平。

三、更加一般化的搜寻技术

在第二部分所得到结论不管多么令人吃惊，它们都依赖于搜寻技术性质的特殊假设。我们假设了尽管各人的搜寻成本是不同的，但是每一次搜寻的成本都与以前进行的搜寻次数不相关。这个关键假设是不可信的。一方面，很多次搜寻的成本是固定的：一旦某个人抵达了某个购物中心，去另外一家商店的边际成本是非常小的。另一方面，在搜寻次数大于某个值之后，搜寻成本会随着搜寻次数的增加而增加：因为时间和金钱变得更加稀缺了。尽管前者的效应使得单一价格水平均衡更加不可能存在，但是后者的效应使得需求曲线会产生如图1所示的弯折，并且价格可能会随着企业数目的增加而增加。当市场上的企业数目众多，消费者很可能会颇费一番工夫，进行搜寻后才能找到价格较低的企业。因此找到价格较低的企业的期望边际成本的增长速度可能会快于 N。这会使得价格降低时的价格弹性降低，从而产生了弯折。而市场之中的企业数目越大，需求弹性就越小，因此就可能会产生市场价格伴随企业数目的增加而上升的情况。

我们将会在两种有限的情况下推导出结论，而不准备建立一般性的结果。

1. 凸搜寻成本函数

假设最多允许同一消费者有两次搜寻行为：

$$\text{对于 } t > 2, s_1 = 0, s_2 = s, s_t = \infty$$

这样消费者在搜寻之后，回到原来的企业是否不会产生搜寻成本，就显得很重要了。为简单起见，假设不会。

没有重置搜寻

ϵ^+的值是不变的，但是位于高价企业的消费者会开始搜寻低价企业的充分必要条件是

$$\frac{v(p)-v(p^*)}{N-1} \geqslant \hat{s}$$

在这些搜寻之中，只有1/（$N-1$）次它找到了低价企业，于是

$$M = L\left[\frac{F(\hat{s})}{N-1}\left(\frac{N-1}{N}\right)+\frac{1}{N}\right]$$

于是

$$\epsilon^- = E(p^*) + \frac{f(0)p^*x(p^*)}{N-1}$$

当N增加，它的值会下降。对于$N>2$的情况，需求曲线会发生如图3的弯折。当$N\to\infty$，价格水平不能确定的范围增加了，$\tilde{p}\leqslant p\leqslant p^m$，价格可以是位于垄断价格$\tilde{p}$中间的任何水平，而$\tilde{p}$的大小由式（15）决定。由于$\epsilon^-$是递减的，很明显随着$N$的增加，市场价格水平是可能上升的。

存在重置搜寻

通过直接计算可以得到：

$$\epsilon^- = E(p^*) - fxp^*\frac{N-1}{N^2}$$

$$\epsilon^+ = E(p^*) + fxp^*\left(\frac{N-1}{N}\right)^2$$

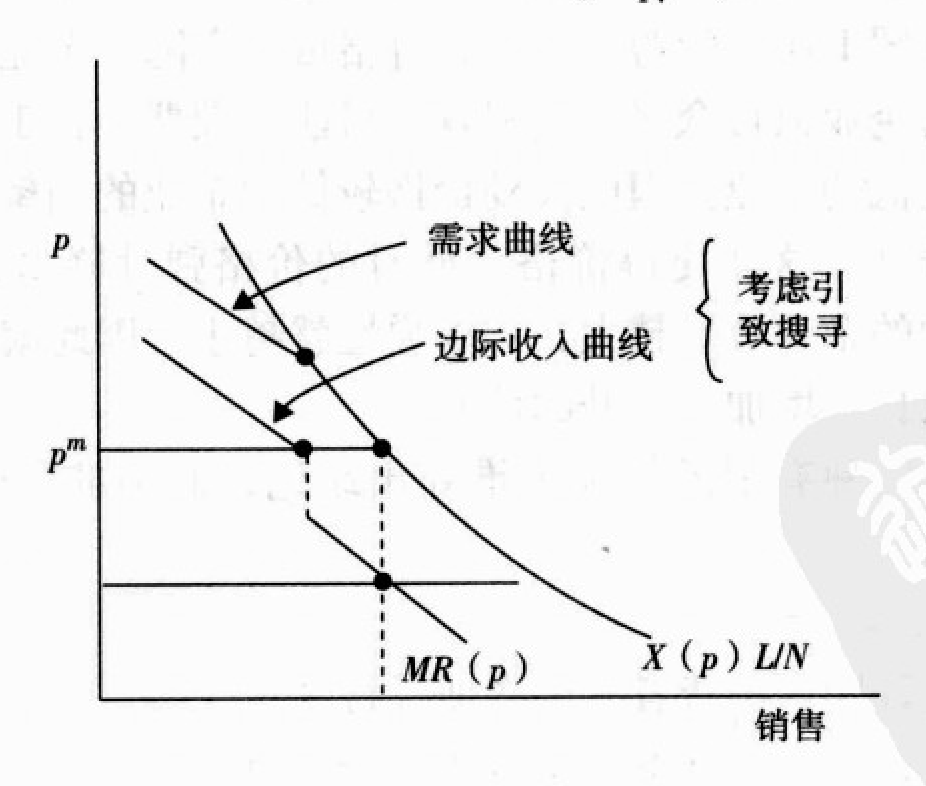


图3　如果f（0）=0，在$p=p^m=p^*$不会有弯折，唯一的均衡为垄断价格均衡

同样地，当 $N>2$ 需求曲线会发生弯折：当 $N>2$ 有 $\epsilon^{+}>\epsilon^{-}$。随着 N 的增加，ϵ^{+} 是增加的而 ϵ^{-} 是递减的，而 p 值位于 $\tilde{p}$ 和 p^{m} 之间，随着 N 趋于无穷大，结果达到极限的情况。

2. 凹搜寻成本函数

假设 $s_1=s_2>0$ 以及当 $t>2$ 时，$s_t=0$。如果位于某个企业的消费者通过搜寻（如果他去不止一家商店购物）可以获取完美的信息。于是是否存在着重置搜寻就不那么重要了；价格上升时的搜寻条件就又是式（8），因此 ϵ^{+} 仍然可以由式（11）描述。

而价格下降时的搜寻条件同样可以由式（8）描述，期望销售额可以由式（12）描述。于是

$$\epsilon^{-}=E(p^{*})+f(0)xp^{*}(N-1)\geqslant\epsilon^{+}$$

在 $N>2$ 的条件下，不等式取严格的不等号：如果存在着多于两个企业，那么市场就不存在单一价格均衡。

四、结果的稳健性

本文证明了已经在文献中有过细致探讨的续贯搜寻模型在消费者存在着严格正的搜寻成本，以及企业数目为连续统的条件下是十分特别的。我曾经提到：那个模型的结果——唯一的均衡使得位于垄断价格水平的单一价格均衡——是不稳定的。相反，我认为我的结果（即均衡价格水平会低于垄断价格水平，并且存在着价格的差异）是稳健的。正是由于缺乏足够大的利润差异，使得在消费者存在着严格正的搜寻成本以及企业数目为连续统的条件下，均衡价格水平位于垄断价格水平。而一般的情况，正如本文所示，是存在着足够大的利润差异的。当存在着足够大的利润差异时，因此市场最高的售价要低于垄断价格水平。比如，当市场之中存在着商品差异，或者个人对于价格分布具有不完美信息，并且（可以想象）每个人所拥有的先验信念是不同的，如果企业选择降价，它们不仅仅会对于现有的消费者出售更多的商品，而且还会发现其实还有更多的愿意购买商品的消费者。

总之，我注意到价格上升时的需求弹性不同于价格下降时的需求弹性。在凹搜寻成本或者线性搜寻成本，并且不存在重置搜寻的条件下，我们得到了均衡时一定存在着价格差异的结论。这与 Salop and Stiglitz（1977）早期得到的结果一致。在凸搜寻成本条件下，我们得到了存在弯折的需求曲线。注意这是在完全竞争条件下得到的存在弯折的需求曲线，而不是在传统的寡头垄断控制下的竞争条件下得到的。

五、结论

令人吃惊的是，尽管搜寻成本的重要性在很久以前就被人们认识到了，但是到目前为止，关于市场均衡性质的真正含义，我们却知之甚少。

本文主要的两个目的是：第一个是想验证在搜寻模型的条件下关于企业数目越大竞争就越有效的标准假设是否正确。结果表明市场企业数目与均衡价格之间的关系并不是我原来所想像的那么简单。传统的结论只是在线性搜寻成本条件下成立。而在凸搜寻成本条件下，随着企业的数目增加，均衡价格的不确定性也越来越大了。尽管均衡价格最小值是降低了，但是均衡价格的最大值却上升了：在极限条件下，均衡价格的最大值等于垄断价格水平。[16]

第二个目的是为了确定在何种条件下，搜寻占主导的市场均衡之中企业会面临弯折的需求曲线，因为这会导致某些宏观经济学通常意义上的刚性。我已经证明了只要搜寻成本是凸的，情况就是如此。比如工资水平的变化的结果可能不会导致任何的调整，无论是价格还是产量。需求曲线的弯折部分的位置取决于个人对于各个企业的信念（或者更加准确地说，由于真正重要的是企业对于其所面临的需求曲线的看法，因此需求曲线的弯折部分的位置，取决于企业对于消费者关于价格分布预期的预期），因此，如果所有个人（企业相信是所有个人）相信所有（其他）企业会提价百分之五，那么在弯折部处的产出价格会上升百分之五。此模型与名义价格和实际价格刚性相一致。这里存在着多重均衡。当然，我并没有给出一个（企业期望的）消费者预期形成的理论，换句话说，某些惯例是如何形成理论的。比如说如果不同企业的投入成本是高度相关的，并且所有企业过去都传导过成本上升，但是没有传导过成本下降，那么如此一来，在弯折处的价格就只会随着投入成本的增加而增加，但是不会随着投入成本的下降而下降。

虽然该理论可能对于消费商品市场比对于其他商品市场更加适用，但是要注意到如果商品是有差异的，所导出的生产者需求曲线本身就会具有弯折，这反映了零售商所面临的带弯折的需求曲线。

16　相比之下，Satterthwaite（1970）得到了一个客观的结果。除了上面提到的模型与脚注 13 标明的 Satterthwaite 模型之间的差异之外，还有一个很重要的区别值得强调。在这个模型之中，显含了企业会考虑到通过降低价格，使那些本来不会进行搜寻的搜寻成本较低的消费者会因此采取搜寻。Satterthwaite 忽视了这个效应。由于当企业数目不大的时候，这种引致搜寻的效应是最大的，所以，我原本认为价格与企业数目可能存在的正相关关系也许要比 Satterthwaite 模型还要大。

本文所建立的模型基本上是静态的。消费者搜寻并且购买消费品的时期只有一期。而在很多情况下，一旦消费者决定在某家企业购买商品，其后的几个时期之内，他都会留在那里。采取降价的措施来驱逐别的企业的“忠诚的”消费者是十分困难的，这意味着价格下降的需求弹性是很低的。这加剧了本文之中分析的能够引致刚性的弯折现象。不幸的是，正式地分析这样一个动态模型，会大大超出这篇短文的篇幅。

附录Ⅰ 多重价格均衡：一个例子

在我的多重均衡例子中，有两个企业的售价较高 p_1，而另外一个企业的售价较低 p_2，而位于高价企业的顾客之中有的 $F(\hat{s})$ 比例的消费者会继续搜寻，其中

$$v(p_2) - v(p_1) = \frac{3\hat{s}}{2}$$

（由于平均要花费消费者 3/2 次搜寻才能找到低价企业，此处假设不存在重置搜寻）。高价企业的销售额为 $[1 - F(\hat{s})]/3$，低价企业的销售额为 $[1 + 2F(\hat{s})]/3$。

如果有一个高价企业提价 Δ，那么它会损失一些消费者。任何离开最高价（要价 $p_1 + \Delta$）企业的消费者会持续搜寻直到找到最低价企业为止。因为以五五开的机会节省 Δ，或者 $p_1 + \Delta - p_2$ 是很划算的，当然如果节省 $p_1 + \Delta - p_2$ 的概率是 1 的话，那么就绝对是有利可图的了（每单位购买的商品）。于是，这样的消费者在 $s < \hat{\hat{s}}$ 的条件下会搜寻，其中

$$\frac{3\hat{\hat{s}}}{2} = v(p_2) - v(p_1 + \Delta)$$

如果企业降低售价，同理可证，任何愿意离开最高价企业以搜寻第二高价企业的个人都会愿意搜寻最低价企业。于是，它的销售额仍然是 $[1 - F(\hat{\hat{s}})]/3$ 其中

$$\frac{3\hat{\hat{s}}}{2} = v(p_2) - v(p_1 + \Delta)$$

这样它的需求弹性是

$$E + \frac{2f(\hat{s})xp}{3[1 - F(\hat{s})]}$$

以及

$$\frac{p_1 - c}{p_1} = 1 \Big/ \left\{E + \frac{2f(\hat{s})x_1p_1}{3[1 - F(\hat{s})]}\right\} \tag{A1}$$

其中 $x_1=x(p_1)$。同样的方法，可以得到低价企业的需求弹性为

$$E+\frac{4f(\hat{s})xp}{3[1+2F(\hat{s})]}$$

令 $x_2\equiv x(p_2)$[17]可以得到

$$\frac{p_2-c}{p_2}=1\bigg/\left\{E+\frac{4f(\hat{s})x_2p_2}{3[1+2F(\hat{s})]}\right\} \tag{A2}$$

假设需求弹性为常数。当 $p_1>p_2$ 有

$$\frac{2x_2p_2}{1+2F(\hat{s})}>\frac{x_1p_1}{1-F(\hat{s})}$$

或者

$$F(\hat{s})<\frac{2(p_2/p_1)^{1-E}-1}{2[1+(p_2/p_1)^{1-E}]} \tag{A3}$$

当 $F(\hat{s})<1/4$ 时，满足 $\frac{2f(s)x_1p_1}{3[1-F(\hat{s})]}\geqslant f(0)x^*p^*$。比较式（14）和式（A1），可以看出，

$$当\frac{2f(s)x_1p_1}{3[1-F(\hat{s})]}\geqslant f(0)x^*p^* 时,有 p_1\leqslant p^* \tag{A4}$$

而由式（14）和式（2）

$$当\frac{4}{3}\frac{f(\hat{s})x_2p_2}{1+2F(\hat{s})}\geqslant f(0)x^*p^* 时,有 p_2\leqslant p^* \tag{A5}$$

例如，如果 $E=1$，$F(\hat{s})<1/4$，并且 $f(0)<8/9f(\hat{s})$，$p_2<p_1<p^*$，两种价格水平都要低于对称的双头垄断情况下的价格水平。很明显，在其他的搜寻成本假设条件下，不等号可能完全相反。同样地，如果 $E=1$，$F(\hat{s})<1/4$，并且 $f(0)<f(\hat{s})$，两种价格水平都要低于对称的三头垄断情况下的价格水平（此处存在着重置搜寻）。

附录Ⅱ　在企业数目为连续统的情况下均衡价格分布

让我们把注意力集中在生产成本为零，并且所有的消费者都只消费一个单位的特殊情况，条件是价格水平要小于 u。如果 $G(p)$ 是价格分布，回

17　很容易构造满足式（A1）和式（A2）的数字例子。因此，如果 $E=0$，$x=1$ 以及 $c=1$，那么就存在着 $p_1=11/8$，$p_2=10/8$，$\hat{s}=2/3\times1/8=1/12$，$f(\hat{s})=18/5$，$F(\hat{s})=1/10$ 的均衡。值得注意的是，在构造这个例子的时候，只有搜寻成本的局部性质发挥了作用，尤其是 $f(\hat{s})$ 以及 $F(\hat{s})$，因此，分析同样适用于 $f(0)=0$ 的情形。

忆搜寻成本为 s 的消费者的保留价格水平由下式给出

$$\hat{p} = \hat{p}(s) = \frac{\int_0^{\hat{p}} p dG(p)}{G(\hat{p})} + \frac{s}{G(\hat{p})} \tag{B1}$$

那些搜寻成本为 s 的消费者会在所有的低价企业之间平均分配他们的需求；即每一个低价企业所得到的顾客都 $f(s)/G(p)$ 成比例。任何企业的销售额都与下式成比例：

$$\int_{\hat{s}(p)}^{\infty} \frac{f(s)}{G(\hat{p}(s))} ds \tag{B2}$$

其中 $\hat{s}(p)$ 是保留价格为 p 的消费者的搜寻成本。这样，我们就要求

$$\pi = p\int_{\hat{s}(p)}^{\infty} \frac{f(s)}{G(\hat{p}(s))} ds = \kappa \text{ 为常数} \tag{B3}$$

对于所有的价格分布都成立，并且对于所有价格水平有

$$\pi = p\int_{\hat{s}(p)}^{\infty} \frac{f(s)}{G(p(s))} ds \leqslant \kappa$$

将式（B1）逐项相加，可以得到

$$\begin{aligned} s &= \hat{p}G(\hat{p}) + \int_0^{\hat{p}} G(p)dp - pG\big|_0^{\hat{p}} \\ &= \int_0^{\hat{p}} G(p)dp \equiv H(\hat{p}) \end{aligned} \tag{B4}$$

并且

$$\frac{ds}{d\hat{p}} = G(\hat{p}) = H' \equiv h(\hat{p}) \tag{B4'}$$

这样我们可以把式（B3）写为

$$\pi = p\int_p^{p_{\max}} f(s(p))dp + [1 - F(s(p_{\max}))]p \tag{B5}$$

并且

$$\begin{aligned} \left.\frac{d\pi}{dp}\right|_{\hat{p}} &= \int_{s(\hat{p})}^{\infty} \frac{f}{G} ds - \frac{\hat{p}f}{G}\frac{d\hat{s}}{dp} \\ &= \frac{\kappa}{\hat{p}} - \hat{p}f = 0 \end{aligned}$$

于是有

$$\hat{p}^2 = \frac{\kappa}{f(\hat{s})} \tag{B6}$$

结合式（B5）可以得到条件

$$1 = F(s(p_{max})) + \frac{\kappa}{p_{max}} \quad (B7)$$

因此由式（B6）可以得到

$$\frac{ds}{dp} = -\frac{2f}{f'p} = G > 0, f' < 0$$

$$\frac{d^2s}{dp^2} = \frac{2f}{f'p^2} + \left(\frac{2}{p} - \frac{2f}{f'^2}\frac{f''}{p}\right)\frac{2f}{f'p}$$

$$= \frac{2f}{f'p^2}\left(3 - \frac{2f}{f'^2}f''\right) = g > 0$$

其中 $g = dG/d(p)$，因此如果均衡价格分布是存在的话，有

$$f' < 0, \frac{ff''}{f'^2} > \frac{3}{2}$$

为了构造满足条件的例子，令

$$\frac{f''f}{f'^2} = m > \frac{3}{2}$$

相加得到

$$f(s) = (b + as)^{1/(1-m)} \quad (B8)$$

再相加得到

$$F(s) = (m-1)\left[\frac{(b+as)^{(m-2)/(m-1)} - b^{(m-2)/(m-1)}}{a(m-2)}\right] \quad (B9)$$

将其替代进式（B6）再反代，得到当 $1 \leqslant p \leqslant u$ 时有

$$s(p) = \frac{(\kappa/p^2)^{1-m} - b}{a}$$

$s(1) = 0$ 意味着

$$b = \kappa^{1-m} \quad (B10)$$

由式（B4′）可以得到：对于 $1 \leqslant p \leqslant u$，有

$$G(p) = \frac{\kappa^{1-m}2(m-1)p^{2m-3}}{a} \quad (B11)$$

这表明如果 $G(u) = 1$，那么

$$a = \kappa^{1-m}2(m-1)u^{2m-3} \quad (B12)$$

重写此式得到

$$G(p) = \left(\frac{p}{u}\right)^{2m-3} \quad (B13)$$

因此在 $p = 1$ 这一点存在着聚点，$G(1) = u^{3-2m}$，条件（B7）就变为（利用前面得到的条件）

$$1 = \frac{\kappa}{2(m-2)}\left(\frac{1}{u} - u^{3-2m}\right) + \frac{\kappa}{u} = t(u) \qquad (B14)$$

由于 t （1） $=\kappa$ 和 t （∞） $=0$（因为 $m>3/2$）对于任何 m（$m>3/2$，$m\neq 2$）以及 k 都存在着满足式（B14）的 u 值。因此如果搜寻强度的参数 a 和 b 能够满足式（B10）和式（B12），那么均衡价格分布就能够构造出来。例如，令 $m=3$，并且 $\kappa=2$，于是

$$f(s) = \frac{2}{(1+4su^3)^{1/2}}$$

这样就存在着价格差异的均衡，其形式为：对于所有 $1\leqslant p\leqslant u$，有

$$G(p) = \left(\frac{p}{u}\right)^3$$

一般而言，分布之中的最高价格是低于垄断价格的；因为如果得到和其他企业相同的利润水平，售价最高的企业必须找到它的顾客，即所有的搜寻成本高于或者等于 $\bar{s}$ 的消费者。但是这里还存在着某个消费群对于是否进行搜寻是无差异的，如此一来，最高售价的企业的需求弹性要大于个人的需求弹性（价格的最大值可能低于也可能高于对称均衡价格）。

参考文献

Arrow, Kenneth J., and Rothschild, Michael, "Preliminary Notes on Equilibrium Price Distributions with Limited Information." Working Paper no. 34. Stanford, Calif.: Stanford Univ., Inst. Math. Studies Soc. Sci., 1973.

Axell, Bo. "Search Market Equilibrium." *Scandinavian J. Econ.* 79, no. 1 (1977): 20–40.

Braverman, Avishay. "Consumer Search and Alternative Market Equilibria." *Rev, Econ. Studies* 47 (April 1980): 487–502.

Braverman, Avishay, and Dixit, Avinash. "Consumer Search and Market Equilibria: A Note." *Rev. Econ. Studies* 48 (October 1981): 657–58.

Butters, Gerard R. "Equilibrium Distributions of Sales and Advertising Prices." *Rev. Econ. Studies* 44 (October 1977): 465–91.

Diamond, Peter A. "A Model of Price Adjustment." *J. Econ. Theory* 3 (June 1971): 156–68.

Pauly, Mark V., and Satterthwaite, Mark A. "The Pricing of Primary Care Physicians' Services: A Test of the Role of Consumer Information." *Bell J. Econ.* 12 (Autumn 1981): 488–506.

Reinganum, Jennifer F. "A Simple Model of Equilibrium Price Dispersion." *J. P. E.* 87 (August 1979): 851–58.

Rosenthal, Robert W. "A Model in Which an Increase in the Number of Sellers Leads to a Higher Price." *Econometrica* 48 (September 1980): 1575–79.

——, "A Dynamic Oligopoly Game with Lags in Demand: More on the Monotonicity of Price in the Number of Sellers." *Internat. Econ. Rev.* 23 (June 1982): 353 – 60.

Rothschild, Michael. "Models of Market Organization with Imperfect Information: A Survey." *J. P. E.* 81 (November/December 1973): 1283 – 1308.

Salop, Steven C. "Information and Monopolistic Competition." *A. E. R. Papers and Proc.* 66 (May 1976): 240 – 45.

——, "Second-best Policies in Imperfect Competition: How Improved Information May Lower Welfare." Discussion Paper no. 11. Philadelphia: Univ. Pennsylvania, Center Study Organization Innovation, August 1979.

Salop, Steven C., and Stiglitz, Joseph E. "Bargains and Ripoffs: A Model of Monopolistically Competitive Price Dispersion." *Rev. Econ. Studies* 44 (October 1977): 493 – 510.

——, "The Theory of Sales: A Simple Model of Equilibrium Price Dispersion with Identical Agents." *A. E. R.* 72 (December 1982): 1121 – 30.

——, "Information, Welfare, and Product Diversity," In *Arrow and the Foundations of the Theory of Economic Policy*, edited by George R. Feiwel. London: Macmillan, 1987.

Satterthwaite, Mark A. "Consumer Information, Equilibrium Industry Price, and the Number of Sellers." *Bell J. Econ.* 10 (Autumn 1979): 483 – 502.

Scitovsky, Tibor. "Ignorance as a Source of Oligopoly Power." *A. E. R. Papers and Proc.* 40 (May 1950): 48 – 53.

Stiglitz, Joseph E. "Equilibrium in Product Markets with Imperfect Information." *A. E. R. Papers and Proc.* 69 (May 1979): 339 – 45.

——, "Equilibrium Wage Distributions." *Econ. J.* 95 (September 1985): 595 – 618.

Stigler, George J. "The Economics of Information." *J. P. E.* 69 (June 1961): 213 – 25.

von zur Muehlen, Peter, "Monopolistic Competition and Sequential Search." *J. Econ. Dynamics and Control* 2 (August 1980): 257 – 81.

信息与规制

信息与规制*

政府参与了许多的规制活动（regulatory activities），例如对公用事业、金融机构的规制，对汽车安全、汽车尾气的排放标准的规定等。规制只是政府用来得到某些期望结果的方法之一。某种程度上，规制是介于更直接控制（Control）的方法（使一个行业国有化）和更间接控制的方法（税收和补贴）之间的。本章的主题是：要理解如何对控制方法进行选择，以及当选择了规制方法，如何设计规制方案这些问题，最核心的就是要理解不完全信息和信息成本导致的一些问题。

本文讨论的顺序安排如下。本文的前一部分总览了与政府控制（government control）有关的一些基本问题。我们的分析按以下内容进行：首先，我们列出了政府控制的原因。其次，我们描述了这里考虑的信息问题的本质，并注意到我们和文献所关注的某些控制问题都可以被看作是更广义上的控制问题的一些特例。我们还列举了在征税、规制和国有企业下不同的控制结构，略述了控制的结构和层级并非不合理的原因，讨论了特许权竞标（franchise bidding）在规制方面所能起到的作用。再次，我们分析了最优税收、规制和垄断定价问题之间存在的明显的对偶，如果不恰当地处理这种对偶，可能导致一些错误的结论。最后，我们略述了从企业中抽租所具有的规

---

* David E. M. Sappington and Joseph E. Stiglitz, "Information and Regulation", *Public Regulation*, E. Bailey (ed.), London: MIT Press, 1987, pp. 3-43. 本文的部分内容是约瑟夫·斯蒂格利茨在胡佛研究所做访问学者时完成的。非常感谢胡佛研究所和国家科学基金的对本研究的资助。本章的观点并不代表 Bellcore 的观点。感谢 Elizabeth Bailey, David Besanko, Steven Salop, Jean Tirole, 以及匿名审稿人所提出的有用建议。

制问题，并由此引出了本文的后一部分。

本文的后一部分着重讨论作为政府控制方法之一的规制。首先对规制者的抽租问题给出了一个正式的陈述。然后给出了不完全信息给规制者带来的问题，并回顾了已被提议的处理这些问题的一些方法。尤其是讨论了包含审查、跨公司比较、跨时间比较的规制方案。我们也尝试着将不完全信息视角下的规制“新”模型和“经典”的规制模型（例如，Averch and Johnson 在1982 年所做的分析）联系起来。

在开始之前，我们要做出一点申明。我们尝试着对有关有限信息下的规制政策设计问题的文献，做一个全面的综述。但是很可能我们没有引用到一些重要的参考文献，而过多地引用了我们最熟悉的文章（例如，我们自己的）。我们提前对这两点导致的不足致歉。我们在开始的时候也注意到了另外两个备受推荐的研究报告，可以作为我们的补充。同我们的讨论最相关的是卡洛德等人（Caillaud et. al.，1985）所做的研究。另一个是哈特和霍姆斯特姆（Hart and Holmstrom，1985）对有关合约的文献所做的综述。

1. 政府控制的原因

就像我们在一开始就注意到的，我们观察到政府控制至少采取三种不同的形式：间接控制（indirect control）、规制（regulation）和国有化（nationalization）。一般是“市场并非总是运作良好”的信念导致了以上三种形式的政府行为。现代经济学的一个中心结论是福利经济学的基本定理。该定理给出了在哪些条件下竞争市场导致的配置是帕累托有效的，以及在哪些条件下帕累托有效的配置是可以通过市场竞争达到的。当那些条件不被满足时，一般而言，市场将不再有效。因此，原则上政府干预是必要的。

大多数的规制是针对三类的市场失灵：

（1）不完全竞争：在多种情况下，市场都不是竞争性的（经济上的效率要求的不仅仅是竞争，而是要求公司都是价格的接受者，但即使市场中存在很多公司，它们也可能并非是价格的接受者）。尤其是当大量的非凸性存在并导致自然垄断时，情况更是如此。这就给政府干预公用事业提供了动机。[1]

1 最近的文献（Baumol，Panzar and Willig，1982）不仅关注了非凸性的重要性，还关注了沉没成本的重要性。在没有沉没成本的时候，即使有自然垄断，也可能达到有效率的结果。另一方面，即使是非常小的沉没成本（这在经济中是非常普遍的），存在很少的非凸性，都有可能达不到竞争性的并有效率的结果。参见 Dasgupta and Stiglitz（1985）。

(2) 不完全信息：福利经济学的基本定理假设信息是完全的。最近格林瓦尔德（Greenwald）和斯蒂格利茨（1986）已经证明了当存在不完全信息时，政府的干预（即使限制政府只能采取征收商品税和补贴的形式）几乎总是必要的，且会引起帕累托改进。不完全信息为两大类的政府规制行为提供了根据：第一，在保险市场上，被保险事件发生的概率会因为保险的出现而增加，为了消除这种影响，保险公司会试图规制被保险者的行为。因为政府或明或暗地参与了大量的保险活动，它会尝试着实施规制来降低被保险事件发生的概率。[2] 其次，建立消费者保护法和劳工保护法的一个前提假设就是个体无法像政府那样很好地保护自己的权益，而这在很大程度上是因为他们所掌握的信息不完全，并且对他们而言，收集信息的成本很高。

(3) 外部性：当存在外部性时，竞争性市场导致的结果是没有效率的。例如，政府针对控制污染所实施的规制，就是为了减轻外部性造成的问题。

对政府干预行为的批评者认为，即使简单的竞争性模型无法得到有效的结果，但更广义地理解竞争过程，会发现是可以达到的。例如，在外部性的情况下，人们有动力聚集在一起来消除该类不效率的结果。这种观点可以被大致地看作是科斯定理。当然，该定理忽略了所有标准的公共品/搭便车问题。而且它忽略了私人商议中的交易成本：事实上，政府可以被看作是建立起来处理这些外部性的机构。最后，该定理忽略了在小团体内部存在的议价问题和不完全信息导致的在议价中存在的不效率。[3]

当信息完全时，规制、国有化和税收/补贴机制这三种方法都是等价的。[4] 为了说明这点，我们假设政府希望实现某种结果。它可以通过税收和补贴的方法，诱导私有企业做政府希望它们做的事情，也可以通过管理机构命令公司做政府希望它们做的事情。此外，政府也可以自己来完成该任务（通过国有化）。事实上，在某种意义上，规制可以简单地视做是强制实施一种非线性的税收/补贴制，当公司不服从管理者的命令时，就会被处罚数额高昂的罚款（罚款的数额足够大，使得被管制的公司总会自愿地服从管

2 然而，请注意，私人保险公司也会类似地试图“管制”被保人的行为，如果他们有权力这样做。

3 即当议价的双方相互之间都存在信息不完全的问题，就会产生一种高成本的自选择过程，使他们能够将自己区分开来。对此进一步的讨论请参见 Farrell（1985）。

4 当存在非凸性时会出现一些问题；一些配置可能并不能在线性税收—补贴机制下形成。然而，一般而言，其可以在非线性税收/补贴机制下形成（参见脚注5）。

理者的意愿）。[5]当信息完全时，不会出现对政府控制或政府控制方式的效率问题的争议。

当信息不完全的时候，规制、国有化和税收/补贴并不是等价的，并且规制这种方法会变得重要。到目前为止的研究都没有清楚地刻画出这几种控制方法的不同之处。我们的目的是，辨识出一些明显的区别。为了说明存在争议的问题，我们将我们的注意力集中于分析对公用事业的规制行为，但这里的分析明显地也适用于其他规制情况。

用信息理论研究方法来讨论这些问题时，一开始总是要试着描述出：在什么时候，是谁了解什么样的信息，并且在各种情况下，谁有权利采取什么样的措施。一个很自然的假设是，参与生产的人（也就是经营公用事业的人）对以下几方面比参加生产以外的人掌握了更多的信息：（1）生产的技术（包括现在的技术和技术提高的可能性）；（2）产品的需求结构；（3）相关要素成本。[6]其他人（政府或管制者）可能也了解一些关于这些变量的信息，但在快速变化的环境下，其掌握的信息往往都是滞后的，可能并不太适用于当前的情况。正是这种有限信息导致了政府所面临的核心问题。政府只能命令公司做一些可行的事，但它可能并不知道哪些事是可行的。即使它知道哪些事是可行的，它希望公司做的事也取决于要素成本、技术和消费者偏好，而这些信息它掌握得没有公司多——例如，如果降低污染的成本是巨大的，政府可能就不会要求公司大幅度地降低污染级别；而如果成本不是很大，政府就希望公司能大幅度地降低污染。但是如果政府必须依靠公司来告诉它成本有多大，那政府如何选择呢？一般而言，它不能指望公司来告诉它真相——例如，有哪个公司的经理会告诉政府真实的保留工资或者时间的机会成本？在本文的第二部分中，我们大部分具体的建模分析都旨在推导出能够诱导公司揭露出其真实信息的方法（虽然要支付成本，并常常会导致行

5　非线性税收/补贴机制是指对公司的支付或从公司得到的支付不是简单地和相关变量成一定比例。因此污染物的线性价格体系是指向污染者收取的费用为污染量的一定比例。而非线性体系的一种是对超过一定量的每单位污染收取固定的费用，或者对一定量以下的每单位污染收取固定的费用，而对超过的每单位污染收取更高的每单位固定费用。管制条款往往规定的在一定水平以下不会罚款，但超过了一定水平就会有无穷大的罚款。当然，真正实施的罚款很少是无穷大的，污染者不仅可以选择忽略管制者，而且可以忽略之后的法院指令。然而，通常这样做而得到的处罚会足够大，以至于公司会选择服从。

6　当然这里忽略了一个事实——公共事业单位内的信息不是由其中的每个人平等地分享的。在组织的内部以及组织和政府之间存在非常重要的控制问题。我们主要关注后一类问题，虽然我们也会简单地讨论一下前一类。

为的扭曲)。

我们所关心的信息问题的出现，是因为它们会对政府试图控制公司并从公司“抽”租造成影响。我们现在开始直接分析控制问题。

2. 控制的等级

控制问题最简单的模型包括两方当事人：建立控制的一方（委托人）和在这些控制下执行任务的另一方（代理人）。例如，最简单的管制模型考虑一个管制者和一家公司之间的相互作用。事实上，就像第二部分的综述所解释的，这种类型的模型是大多数文献所关注的。然而，在现实中，控制问题要复杂得多。

用管制的情况举例。为了简化分析，我们假设经济体中只有一家公司，该公司生产社会消费的唯一产品。因此在只有一个市场的情形下，我们不用考虑局部均衡和一般均衡分析的区别。即使是这种简单的情况，也有很多的控制问题错综复杂地缠绕在一起。

在传统的 19 世纪的模型中，这样的公司不存在所有权和管理权的分离：所有者—管理者采取任何能够最大化（期望）效用的措施，因为人们很少关注努力的负效用，所以常常假设所有者—管理者是为了最大化其（期望）利润或者公司的（市场）价值。

现在，人们意识到这种方法没有很好地刻画大多数现代企业的情况，这些企业所有权被很多人所有。早期的朴素模型只简单地用“董事会”替代了“所有者”，并忽略了所有者（或董事会）在控制管理者时所面临的任何问题。但董事会的成员（除了那些本身就属于管理层的）往往对管理层所做的工作和公司所面临的所有机会了解有限——他们掌握的有关技术、市场和要素价格的信息有限。这种信息上的限制给了管理层很大的自主权。[7]

但是，即使是强调所有者管理的公司和非所有者管理公司之间区别的观点，也没有能很好地刻画由很多人持股的公司的情况。股东甚至不一定能够有效地行使他们手中由所有权带来的有限的控制权。如同我反复强调的那样，好的管理是一种公共品：公司经营得更好会使所有的股东受益。然而，要了解现在的管理层是否很好地履行了其职责（并试图替换不合格的管理

7 因此有关信息的文献（包括有关委托—代理的文献，Ross，1983；Stiglitz，1974）十分推崇公司的“管理”理论，该理论之前就曾被 Berle and Means（1932）和 March and Simon（1985），以及其他一些学者讨论过。

层）是需要支付成本的。因此每个小股东都有“搭便车”的动机，希望依靠其他股东去做这些管理方面的努力。[8]

然而，并非只有股东受公司行为的影响。债权人不能分享公司的巨额利润，但当公司出现巨额亏损时，会蒙受很大的损失。让一家银行作为牵头银行（lead bank）的制度结构会减轻“搭便车”问题，它在任何情况下都比股东制面临更少的“搭便车”问题，这是因为每个银行都很可能持有某家公司很大的头寸。银行停止贷款的威胁和在规定的情况下，银行有干预公司的特权，使得银行可以对管理层实施一些间接的控制。这就说明应该把公司视作是多个委托人的代理人。[9] 就像现在众所周知的，多委托人—代理人问题的均衡解不是帕累托有效的（Braverman and Stiglitz，1982；Bernheim and Whinston，1984；Arnott and Stiglitz，1985）。图 1 展示了一种最简单的公司内部控制结构。

图 1 也展示了在被管制的情况下的另一个控制问题。当引入规制时，就有另一个“委托人”对公司的活动有一定的控制权。然而，管制者控制的有效性由于信息的有限性而降低了。例如，被规制的公用事业单位可以宣称在目前的回报率下，扩张生产量是不赚钱的，它可以试图制造出一个“危机”。规制者可能很难确认危机是否是真实的。

在被规制的情况下，控制结构还有另一个很重要的问题。一般而言，规制者要么是被选举出来的，要么是由被选举出的官员任命的。为了确保该管制者是出于选民的利益而行事，它可能要服从多种的控制。因此，在实际中，规制者在更多的情况下既是委托人（对公司而言）又是代理人（对其选民或“政府”而言）。我们也注意到司法部往往可以不受约束地对管制者和公司的活动进行检查。

图 12 展示了对国有企业的控制的基本等级结构。对国有企业的经理而言，为下属设计激励机制仍然很重要。此外，司法部仍可以控制企业的活动。国有化带来的不同是董事会不再由股东选举，而是由政府任命的。最初，这点被看作是意味着不需要再设定单独的管制机构对公司实施管制，因为一般都假设公司会按照政府的期望行事，或者董事会能够有效地对公司实

8　因此所有主要的所谓股东控制机制的效果都有限。虽然人们早就意识到了股东大会的无效性，但同收购机制（takeover mechanism）有关的问题最近才引起了人们的关注。例如，参见 Grossman and Hart（1980，1981），以及 Stiglitz（1972，1982，1985）。

9　还存在其他一些受公司影响的当事人：当迁移成本（包括特殊人力资本）为正时，工人就会受到公司状况的影响。

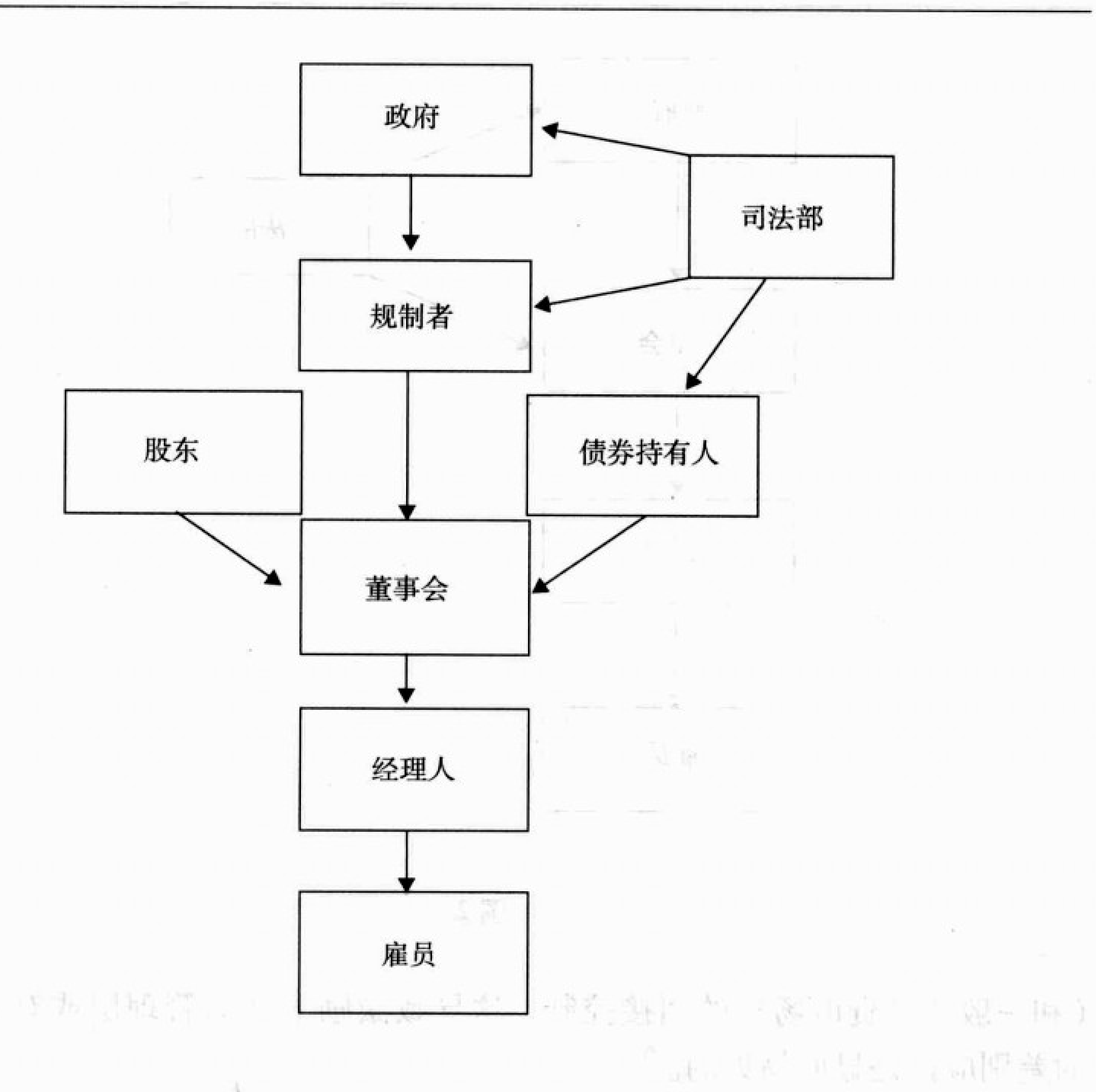


图 1

施控制。

最近，大家的讨论开始强调国有企业和非国有企业（包括被规制的）之间的相似处：公司的行为仍然是由管理者控制的，而管理者有很大的自主权。在每种情况下，董事们都面临如下的激励问题：如何诱使管理者按董事们的所期望的行事，并让他们揭示自己拥有的私人信息，以使得董事们可以辨识应该采取哪些管理措施。

但这还不是唯一的问题。公司的董事往往会认同公司存在的问题，因此把他们自己看作是公司在政府中的代表，代表国有企业的工人和其他成员的利益（例如，争取更大的政府补贴）。因此，至少在一个有很多国有企业的国家，现在关注的焦点是将国有企业“社会化”，以试图改变公司的行为，使公司追求国家目标而不是私人目标。

此外，获得巨额政府补贴的可能性由行政机构的间接控制替代了银行

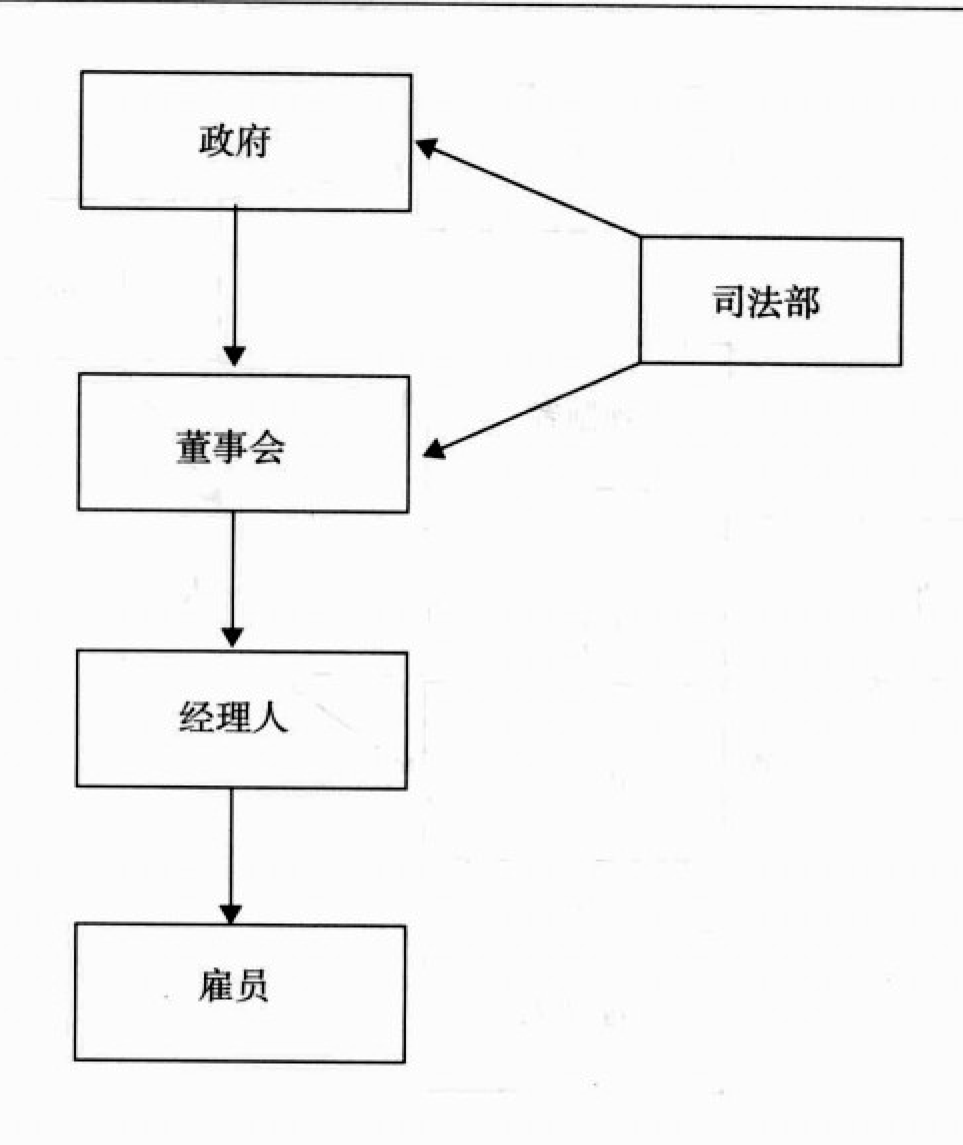


图 2

（和一般的借贷市场）的间接控制。这导致激励上（对管理层或对董事会）的差别应该是显而易见的。[10]

3. 为什么控制是重要的？

在我们的讨论中，我们一直在强调不同制度结构下的控制结构的不同。[11]乍一看，并不能明显地看出为什么公司行为的差别一定和制度结构的不同有关。例如，如果可以在合同中考虑到所有可能发生的事件，则垂直整合对上游企业或者下游企业的行为没有任何的影响。即如我们所讨论过的，同在完全信息的情况下管制和间接控制具有等价性一样，垂直整合了的企业和没有垂直整合的企业也有类似的等价性，即使存在某些种类的信息不完

10 一些讨论已经强调了政府企业的预算“软”约束和私人企业的“硬”约束之间的差别。两者确实存在很多的区别，但不应该过度强调这些区别，尤其是在管制的情况下。企业可以尝试说服政府增加补贴，同样它也可以“说服”管制者提高价格。此外，当破产的威胁逼近时，私人公司常常也试图说服政府给它们提供补贴。

11 传统的经济学理论假设所有的个体都不会犯错，任何表面上看起来不正常的行为都简单地归因于不恰当的激励。相反，最近的研究开始关注过失（其被视作是不可避免的）对组织设计造成的影响。例如，参见 Sah and Stiglitz（1985）。

全。只要可观察的变量集保持不变，无论在哪种制度机构下，都可以为每个参与者设计一样的合同集。而在相同的合同集下，我们可以推测每个参与者的行为也是一样的。

但合同不可能是完全的。因此借贷合同允许债权人可以在特定的条件下以特定的方式干预，但它没有准确地规定债权人可以做什么——即他将采取什么样的直接或间接控制机制（不能得到完全合同的原因在此应该是显而易见的。具体的讨论请参见 Williamson，1975）。因为合同是不完全的，不同制度结构的不同之处在于哪个当事人拥有“剩余控制权”，也就是谁可以采取什么样的应对措施，以及在什么情况下，谁有权干预其他人的活动。并且由于合同的不完全性，有干预权的委托人无法事前承诺自己不采取干预行为。

然而，对为什么控制是重要的疑问仍然存在。在传统的竞争性（阿罗—德布鲁）模型下，控制是不重要的，公司的管理者只是最大化公司的市场价值。任一其他的管理者也会和公司现任管理者采取一样的措施。甚至管理者对任一自然状态发生的可能性的信念对公司的行为没有影响。

在阿罗—德布鲁经济体中，所有的股东对公司应如何行事的意见是一致的。然而，当不存在一套完整的阿罗—德布鲁证券时，对公司是否应最大化在股票市场上的价值，以及任何特定的做法会导致什么样的结果（例如，未来的市场价值），股东间就有可能出现分歧。[12]虽然这些分歧在理论上都具有一定的重要性，但债务人和债权人之间的利益分歧无疑更加重要。尤其是，公司的股东只关心公司没破产时的回报，而公司的债务人只关心公司破产的可能性和破产时的回报（这关系到可以索回些什么）。这时就有真正的

12 斯蒂格利茨（1970）曾讨论过这些问题，部分内容随后发表在其 1972 年的文章中。在讨论中，他注意到在简单的平均方差模型中，股东并不希望公司市场价值最大化，除非他们计划卖出所有的股份。如果股东出售股份的计划不同，他们就不会达成一致。他还证明了如果股东既不打算买也不打算卖股份，则他们会对形成帕累托有效配置达成一致。随后大量的文献（例如，Ekern - Wilson，1974）将后一个结论扩展到许多“生成”条件下，并证明了对价值最大化条件达成一致所需要满足的条件比在存在完整的一套阿罗—德布鲁证券时所需要的条件要稍微弱一些。但如 Grossman and Stiglitz（1980）以及 Radner（1974）所指出的那样，如果公司需要选择负债权益比，则需要存在完整的一套阿罗—德布鲁证券。而且就像斯蒂格利茨最先指出的，以及格罗斯曼和斯蒂格利茨所强调的，在任一存在贸易（或者是因为生命周期的考虑，或者因为不同个体之间获得的信息不同）的经济中，所有个体都打算既不卖也不买证券的假设似乎是不太实际的。

斯蒂格利茨（1972）讨论了当不存在完整的一套阿罗—德布鲁证券时，对公司行为抱有的信念的差别就会很重要。

利益冲突，当破产的可能性显得突出时，这种利益冲突就变得非常重要以致债务人不再信任管理者：他们希望干预公司，而债务合同给了他们这样做的权利。

对于国有企业来说，政府是处于债权人的地位的。政府或者损失投资的回报，或者“被迫”向企业提供更多的资金，认识到这点是非常重要的（在被规制企业中，问题会更复杂，股东可能不被允许保留从“好”决策中获得的利润，而纳税人可能需要“被迫”为“坏”决策买单）。

但同债权人相比，政府会更多地参与公司决策。之前我们讨论过只有在阿罗—德布鲁模型的非常理想化的条件下，公司的目标才会和帕累托有效是完全一致的。因此，原则上，政府应该愿意通过干预来减少公司的一些重要的、可观察到的无效率行为（例如，公司没有考虑到关闭工厂会给政府带来的失业保险方面的成本）。实际上，政府同样也可能因为分配方面的原因而对公司进行干预（例如，为了确保一些偏远地区能够以补助价格获得服务）。在一些情形下，很难区分这两方面的原因（在很大程度上是因为信息方面的原因——这是本章的中心内容）。政府试图降低某一类消费者支付的价格，是为了限制垄断力量，还是为了补贴这一类的消费者呢？

4. 控制方法之一——拍卖

正是由于区分这些分配问题与更狭义上的效率问题的困难，使得一些经济学家认为应该更多地限制政府的作用：增进效率的干预所带来的收益，可能完全被降低效率的干预所带来的损失抵消掉了。

我们暂时假设政府只关心效率问题，并且像常常争论的那样，把分配问题看作是税收范围内的问题。我们还假设政府和可能的生产者都拥有关于消费者偏好的完全信息。为了简化分析，我们考虑以下情况：管制者的目标是最大化某种明确的消费者福利标准，例如，消费者剩余。因此我们去除了任何同控制管制者有关的问题（类似地，我们将去除公司内部的控制问题）。最后，假设存在很多公司，这些公司都拥有或能够获得技术信息，但管制者无法获得这类信息。

在这种情况下，拍卖被提议用来弥补规制者的有限信息。例如，勒布和马盖特（Loeb and Magat，1979）建议规制者可以授予一种垄断特权并根据垄断者产生的消费者剩余的大小来补贴生产。补贴确保了按边际成本定价和最少的生产成本。并且，当竞标的竞争充分时，政府可以占有所有的垄断租。这样，规制者就可以实现当它拥有关于更有效技术的完全信息时所能实

现的结果。

当然，要让这种机制按照以上分析中所述的那样良好运作，必须满足许多条件。尤其是这个理论中暗含了许多的制度假设。例如，假设不受限制的补贴是可行的（并且是无成本的）。如果补贴是不可行的（因为政治方面或其他方面的原因），使得公司唯一的收入必须来自于它的销售收入，则除非公司能够实行完全的价格歧视，否则不能实现有效定价。并且，当各家公司在实施价格歧视方面的能力不同时，一家公司愿意为垄断特权出价更高，可能并不是因为它是一个更有效率的生产者，而是因为它有更好的办法实施价格歧视。

对这个被提议的拍卖机制而言，竞争充分的投标才能够消除所有的垄断租。实际上，公司准备参与投标是有成本的（而这决定了垄断特权所能带来的利润的期望贴现值）。并且由于这些成本一般都是沉没成本，只有少数的公司可能对垄断特权投标。[13]

因此，虽然技术信息由多家公司掌握，可以有助于减轻管制者信息的有限性导致的控制问题，但拍卖不一定是一剂“万能药”。

在一些情况下，即使缺乏完全价格歧视的能力，拍卖也将产生有效的结果。假设所有的个体都是同质的，并且只有线性价格（也就是每单位消费品的价格都是常数）是可行的。那么，如果政府对消费者效用水平设定了下限，并且该下限设定得合适，那么，拍卖就会产生帕累托有效的结果。高于边际成本的涨价不会是统一的，而是将对应于拉姆齐定价（Ramsey price），弗兰克·拉姆齐（Frank Ramsey）在 1927 年讨论了这点。[14]在独立需求曲线的简单情况下，需求无弹性的商品会相应有较高的涨价幅度。

这个结论来自于垄断者的最优定价问题和最优税收问题之间的对偶：后者在利润约束下最大化效用，而前者在效用约束下最大化利润。这两个问题的拉格朗日函数的形式一样，而当约束条件的取值适当时，会得到相同的解。但这个结论需要一些关键假设，我们很快会回到这个问题上来。

---

13 在本文后面，我们将讨论政府如何加强拍卖中的有效竞争。该讨论是对 Williamson（1976）分析过的拍卖的许多制度性缺陷的概述。

14 Boiteaux（1956）是第一个研究公共事业单位定价问题的。在过去的 10 年里，出现了大量的文献在讨论这个问题。研究税收问题，请参见 Atkinson and Stiglitz（1980）；关于最优定价问题，请参见 Baumol and Bradford（1970）。

5. 控制点

到此为止，大部分的分析都是局部均衡分析。但我们也应该考虑一般均衡分析，因为一般均衡分析可以将控制问题进一步复杂化。为了说明这点，假设经济体中有许多被管制的市场，每个市场都生产一种产品。我们还假设每个市场的管制者知道自己的市场中消费者的需求是如何随着产品的价格变化的，但是管制者没有关于整体需求的信息。我们另外假设管制者的目标是最大化消费者剩余。类似地，假设每种产品都有许多可能的生产者，他们都完全了解生产该种产品的成本，但没有一个掌握了所有的技术信息。为了简化分析，我们还假设每个部门中的公司的价格歧视能力都是一样的。

由前面的讨论立即可知，如果允许每个市场的监管者补贴它管辖的生产者——如果前面提到的另外一些市场失灵的情况都没有出现——它可以通过拍卖在市场上实现它最希望得到的结果。即使没有补贴，如果可能的生产者能够和消费者签订有法律约束力的契约，则也可确保出现最被期望的价格向量。通过将特权授予签署合同的能产生最多消费者剩余的公司，就可以达到这一目的。竞争将消除任何垄断者的租，并确保每个市场上成本最低的生产者会赢得特权。[15]

如果补贴行为不被允许，则每个市场上的价格将不等于边际成本。显然，这时的解不如当补贴可行时的解。但在给定约束下，这个结果同在采取直接控制、有关技术可能性的信息完全的体系下的结果是一样的。

如果每个市场都有多个生产者，但不允许跨市场的补贴，则会产生更复杂的问题。盈亏平衡所要求的价格是否是帕累托有效的？这个问题看起来和前面提出的问题很相似，前面我们讨论到最优税收问题和垄断定价问题之间的对偶确保了定价结构的效率。但是当时我们忽略了被监管市场和经济体中其他部分的相互作用，而在政府直接控制时，是会考虑到这种相互关系的。我们将证明在一般均衡分析中呈现在眼前的，绝不是这种对偶结果。政府问题的全局最优将考虑每个公司的定价决策对以下两方面的影响：（1）经济中在别处产生的税收收入（和利润）；（2）其他商品的价格。而前面给出的分权的、局部均衡的方法没有考虑到这些方面。

---

15 并且如果我们对垄断者的能力有完全的信心，我们就应该相信他会利用广义上的特权范围内所有的更有效率的创新。但是，如果公司内部存在之前讨论过的信息问题，则我们会关心组织懈怠的问题，因此我们也没有完全的信心相信所有的更有效率的创新都会被采用。

因此，如果煤炭被用来发电，铁路垄断者会忽略其定价策略对公用事业单位为了实现盈亏平衡的定价的影响。因为该部门的价格超过了边际成本，这种明显的金钱外部性（pecuniary externality）会对福利产生真正的影响（Ebrill and Slutsky，1985；McFarland，1985）。类似地，电价的变化会影响到对石油的需求，从而影响到政府从横财税（windfall profits tax）中获得的财政收入。

管制的分权法还存在一个问题。单个产品的最低成本生产者不一定是两个或更多产品的最低成本生产者。因此当特权的授予是取决于生产者在某个市场上的表现时，全局最有效率的生产者就很可能被忽略，尤其是当各个特权不是同时被授予的时候。

在实际中，垄断者被授予的常常是一种有限的垄断权。因此，考虑到竞争者的进入（生产“相关”产品），垄断者会选择自己所能想到能够阻碍进入者的价格。为了说明这点，假设存在一些生产某种商品的“低效率”的技术，与之关联的成本对垄断者能够制定的价格形成了约束。如果对该产品的需求是非常无弹性的，则垄断者的定价将不得不低于无约束的拉姆齐价格。如果能独立运行的技术是报酬递减的，垄断者在计算“拉姆齐”价格时，将只考虑他的剩余需求曲线（考虑了其他人提供的数量）。因此，这些价格可能会明显地不同于政府希望制定的价格。

当然，如果“恰当地”定义垄断特权，就可以避免所有的这些问题。尤其是如果特权是针对整个经济体的，则局部均衡方法和一般均衡方法之间的区别就消失了。但是要想提议的竞标机制运作良好，需要许多公司都拥有详细的全部技术信息。显然，这不太可能。这个机制要想高效地运作，政府也需要知道全部的需求函数，这样的信息也是不可能获得的。事实上，授权给每个市场中的每个管制者的原因就是要他们形成专门的技术，并用该技术做出正确的判断。管制者获取和处理信息的能力都不足，不能指望他们能够形成全面的专业技术。

因此，一定程度的分权是不可避免的。但我们已经证明了，即使对局部的需求和成本结构的信息是完全的，管理者一般也无法实现全局最优的结果。

我们的分析已经说明了恰当地定义特权是非常重要的。即使我们只考虑

静态的市场进入阻挠问题和剩余需求问题，这也并非易事。[16]在动态的情况下，授予一个广义的特权有可能会在今天形成“正确”的价格。但广义的特权意味着发展出来的一些新的、更有效的技术可能并不会被采用。

当存在异质的消费者，而且政府或垄断公司并不完全知道他们的特征时，标准最优税收问题和标准垄断定价问题之间的对偶将因为另外一个原因而不再成立。在这样的情况下，政府可能会以公平的方式，采用非线性定价标准来分配效用的固定成本负担。在设定价格结构时，政府或许知道一些个体比另一些个体更有支付能力，但只能通过如消费电量一类的指标来区分他们。帕累托有效的歧视性价格结构是指在其他组的消费者保持一定效用水平的约束条件下，能最大化一组消费者的效用，并能确保任一组消费者都不会消费为其他消费者设计的（价格—数量）商品束。后一个约束条件被称做是自选择约束。[17]

而在这种情况下，垄断者的最优定价问题也不完全等价于帕累托有效税收的问题，因为垄断者仅关心在自选择约束下最大化其收入。特别是，对边际以下的消费者（也就是能够从消费商品中获得严格正的消费者剩余的消费者），垄断者只关注定价政策的改变会如何影响到需求和收入，而不是如何影响到消费者的效用。[18]此外，在一些重要的情况下，政府偏好于混合的结果（消费者的特征不能通过他们的消费模式来辨别），虽然垄断者总是偏好于区别对待消费者。[19]

因此，很明显，政府定价问题和垄断者定价问题在形式上的相似仅仅就是一种形式上的类似。政府一般不会简单地希望把定价问题授权给垄断者来处理，即使政府可以通过拍卖的方法获得租。我们的分析已经证明了

16 这里要注意，我们忽略了这样的事实：在实际中，拉姆齐定价常常被误用，只关注结论——应该对需求无弹性的商品征收更重的税。这会忽略：（1）潜在的重要的交叉弹性效应；（2）重要的分配上的考虑常常要求对有需求弹性的商品征收更重的税（Atkinson - Stiglitz 1972, 1980）。

17 当然政府可能并不进行价格歧视。事实上，如果可以选择一个最优的收入税用来再分配收入，以达到最优的目的，那么在某些可分离性条件下，如果消费产出带来的消费者剩余足够高，随边际成本一起收取一个固定数额的进入费是最优策略。进行任何的再分配都最好通过税收体系进行。

18 严格地讲，该结果可能仍在帕累托有效结果集中，只是所有个体的效用的权重为零。但这并不对应于任何“合理的”社会福利函数。

19 在保险的例子中这点似乎是最明显的。如果每个人的风险都是一样的，但风险规避程度不同，则垄断者会实施价格歧视，但一个坚持平等主义的政府会以精算出的公平赔付率为每个人提供完全保险。

垄断者不会像政府一样行事，如果政府像垄断者一样拥有关于技术和需求的信息。

实际上，政府面临的问题是它既不了解技术也不了解需求结构。因此它会发现很难确认边际成本的大小，因此也很难确认任一特定服务的隐含税收。而政府在决定任一特定的歧视性定价机制所带来的净损失时也会有困难。我们知道垄断者不会为了公众利益而行事，但我们现在知道如何采取适当的纠正措施。

6. 从已建立的公司抽租

这些重要的信息问题虽然在政治争论中非常突出，但最近的有关规制问题的文献却没有将其作为中心问题来讨论。最近关于管制的文献更多地关注了以下问题：在局部均衡的情况下，已有公司存在，使得政府无法对特权进行“公平”拍卖，此时政府如何从公用事业单位中抽租。如果政府试图从公司里抽太多的租，公司可能会停止经营。政府必须推断出关于技术的信息，但政府只能用通过设定公司的选择来诱使公司揭示真实特征的办法来获取信息，所以公司的行为被扭曲了。本文的下部分更正式地描述了如何实现这点，并找出了一些重要的情形，在这些情形下扭曲的程度受到了限制。

在介绍这些结论之前，我们先给出几个一般性的观察。首先，这些文献关注的是抽租问题，而不是激励问题——给管理者提供恰当的激励使其努力工作。[20]它们关注的是同技术有关的信息的不完全性，而不是同要素有关的。事实上，这些信息问题都是相关的：管制者发现很难知道一些要素（例如，“管理”）获得的回报是否高于其机会成本。因此，观察到的低效率有可能来自不胜任的（但获得过多报酬）管理者（虽然他们努力工作了），而不是来自于有能力的管理者的怠工（的确，作出好的管理决策常常并不比作出糟糕的管理决策难。因为作出更好的管理决策所需要的努力导致的成本并不是由管理者自己承担的（例如，他可以签约雇人来做一些研究），所以很难确认努力导致的负效用的重要性）。

其次，我们将明显地看出规制政策的结构在很大程度上取决于规制者所知道的信息，它所能监管的内容，以及它了解信息的速度。在将要介绍的大多数模型所处的静态环境中，政府可能最终会了解到大部分和技术相关的信息，因此信息不对称的现象会消失。但在真实的世界中，技术（以及需求

20 稍后会关注一些例外的情况。

和要素价格）总是在不断变化的。因此信息不对称问题的严重程度和环境变化的速度以及规制者了解信息的速度有关。

最后，我们再看看“信息揭示问题”——诱使被管制公司通过它所做出的选择来揭示自己的信息。规制者也可以通过稽核，也就是通过直接监管被管制公司的行为或绩效来获得大量的信息。最优监管行为的设计将是另一篇综述的主题。[21]

7. 关于规制者—公司相互作用的一些简单模型

从这一部分开始，我们关注近年来信息经济学方面的文献是如何讨论管制者和被管制公司之间直接的相互作用的。我们尝试着给出一个一般性的框架，使得我们可以用其分析所有已做的研究。我们会讨论文献中各种模型之间重要的相似之处和不同之处，也讨论一些和文献没有太大联系的重要管制模型（例如，Averch-Johnson 1962 年的模型）。最后，我们简单讨论一下文献中相对关注较少一些重要问题。虽然我们正式的文献综述偏重于关注公用事业管制问题的模型，但其中的基本思想的应用是非常广泛的。读者还应该注意到这里的分析所关注的只是整个控制等级的一部分，即关注的是规制者和公司之间的直接相互作用（图 1）。而且“公司”被看作是一个“黑匣子”，还没有谈及公司内部的分类控制问题。

我们将在这里讨论的模型有如下一些主要特征：第一，规制者和公司之间的信息不对称的性质至关重要。如我们前面所提到的，在不同的时点，每个当事人所了解的信息在很大程度上决定了最优管制政策的性质。第二，规制者可使用的政策工具以及使用这些工具的成本都极为重要。对公司必须承受的处罚的限制和对公司活动进行稽核的高额成本（成本可能无穷大的），会严重地限制管制者影响公司决策的能力。第三，我们将证明，规制者和公司之间的相互作用是否是重复性的，会在制定最优规制政策时起到主要作用。第四，我们将考虑管制者的策略是如何随着它所控制的公司的数量的变

21 因此管制者可能将其审查（例如为了确认糟糕的结果是因为努力不够造成的，还是因为运气太差造成的）建立在某种可观察结果之上，例如，产出。所以这使得审查的概率是距离上次审查的时间的增函数，是自上次审查以后的全部产出的减函数（很明显地，如果公司确切地知道什么时候会有审查，它在审查时的行为就有所变动，因此，为了使审查有效，必须随机进行审查）。

在对审查行为的完整分析中，我们必须还要讨论对审查者的激励问题。因此，如果在 t 时刻，一次银行审查发现有一笔被拖欠贷款，而其在前一次审查中就应该被通告，这次的审查者是否和前一次的审查者是同一个人决定了揭露该信息的动机是不同的。另一方面，如果每次派不同的人审查，则审查的成本可能会大大增加。

化而变化的。

8. 一般性模型

首先，我们给出一个简单又非常一般性的模型，其使得我们可以具体讨论这四个特征中的前两个——信息不对称的性质和管制者可用的政策工具集。因此该模型假设存在一个管制者和一家公司，两者间的相互作用是非重复的。稍后，我们会讨论重复博弈和多家公司的情况。

大多数委托—代理分析的一个共同点是委托人（这里是管制者）在同代理人（这里是被管制公司）博弈时，扮演的是斯塔尔伯格（Stackleberg）领导者的角色。也就是说，管制者在任何其他行动发生以前制定一个激励机制的具体条款。在宣布了公司获得的补偿和管制者的行动是如何取决于一些可观察的变量以后，公司，接着是管制者，轮流采取行动，这些行动将影响管制的最终结果。一些模型让每个当事人只行动一次，而在另一些模型中，当事人轮流行动是很重要的，规制者和公司都至少会采取两个步骤。我们这里构造的一般性模型会考虑这两种可能性。

公司初始的行动（a_1^F）常常只是同规制者交流一下有关自己的信息（I^F）。在这次交流同自己拥有的信息（I^R）的基础上，规制者根据宣布了的管制政策的条款来采取行动（a_1^R）。[22]这次行动可能包括稽核公司提供的信息或是决定是否允许公司继续经营。假设规制者和公司在初始行动以后都观察到一个结果（Z_1）。该结果包括公司给规制者的信息和/或规制者稽核该报告的结果。

然后，公司采取第二次的行动（a_2^F），例如这次行动可能是生产或定价决策。接着规制者采取第二次行动（a_2^F），这次行动可能是观察公司的绩效、审核公司的活动，以及/或者对公司进行支付。在这些行动以后，大家会观察到第二次的结果（Z_2）。这次结果可能包括实现了的价格和产出，以及/或者对公司活动审核的结果（例如，实际的成本）。注意，所有的行动、结果、信息集等都可能是向量。

规制政策（R）由规制者的行动和支付机制 $\tilde{P}$（·）构成。该机制规定对公司的支付是可观察变量 Z_1、Z_2、a_1^F 和 a_2^R 的函数。例如，公司可获得的补偿取决于它与规制者的交流、它的绩效，以及规制者的行动。假设管制者

22 这里我们假设管制者会选择自身的行动（假设所有的当事人都能观察到它的行动）并按照它所公布的管制政策中所承诺的方式来补偿公司。在结束前，我们将简单地讨论一下管制者可能无法做出这样的有效承诺的情况。

选择在初始信息下能够最大化期望效用$\hat{E}\{U^R(\cdot) \mid I^R\}$的规制政策。在这里以及之后的分析中，$E$表示对未知变量求期望。$U^R(\cdot)$为管制者的效用函数。该函数包括对公司的支付，规制者的行动（a_1^R，a_2^R），公司的效用水平（U^F），以及规制的结果（向量）（X）。一般将管制者看作是最大化期望消费者剩余，或者生产者剩余和消费者剩余的加权平均。[23]一般而言，公司将最大化自己的期望利润。

因此规制者问题（RP）的一个非常一般化的表达为：

$$\underset{R}{\text{Maximize}}E\{U^R(X,P,a_1^R,a_2^R,U^F) \mid I^R\}(RP)$$

满足约束条件：

$$a_1^F \in \arg\max_a E\{U^F(P,a,a_2^F) \mid H_1^F\}(AS_1)$$

$$a_2^R \in \arg\max_a E\{U^F(P,a_1^F,a) \mid H_2^F\}(AS_2)$$

$$E\{U^F(P,a_1^F,a_2^F) \mid H_1^F\} \geqslant \overline{U}^F, i = 1,2(R)$$

$$P \in \tilde{P}(FP)$$

$$a_i^s \in A_i^s, i = 1,2, s = R,F,(FA_1^s)$$

此处：

$$P = \tilde{P}(Z_1,Z_2)$$

$$Z_1 = \tilde{Z}_1(a_1^F,a_1^R,I^F,I^R)$$

$$Z_2 = \tilde{Z}_2(a^F,a^R,I^F,I^R)$$

$$X = \tilde{X}(a^F,a^R,I^F,I^R)$$

在对（RP）的陈述中，H_i^F，$i=1$，2，表示公司在采取行动a_i^F时所知道的之前的结果。例如，$H_2^F = \{I^F, R, a_1^F, Z_1\}$。而$\overline{U}^F$是公司的保留期望效用水平。

（RP）中的行动—选择（AS）约束表示，在给定的所有可及信息下，公司选择能够最大化期望效用的行动。个人理性约束（R）确保了R向公司许诺，会让其至少获得保留水平的期望效用，这样使公司不会选择把其资源用在别处。

（RP）中的第四个约束定义了可行的支付机制（FP）。例如通过恰当地规定$\tilde{P}$——可行支付机制集，可以设定奖励的上限和惩罚的下限（例如，破产限制）。（FA_i^s）约束简单地定义了管制者（$s=R$）和公司（$s=F$）的可

23 在实际中，管制者也同样可能想要最大化他们自己的期望效用（例如，收入、任期和声誉），在他们的法定委托权的限制之下。例如，斯蒂格利茨（1971）就注意到了这种可能性。

行行动集。例如，这些约束可能规定了公司和管制者之间的交流，并且/或者规定了管制者不能剥夺公司的财产。

9. 特例：逆向选择和道德风险

文献中的许多管制模型都是（RP）的一些特例。这些模型的区别仅在于对行动的解释，对支付机制的限制，以及信息的性质和其被管制者和公司获得的时间。在详细描述这些模型之前，让我们重申一下，只有当公司的私人信息多于管制者，并且/或者管制者无法无成本地观察公司的行动时，管制者才会面临控制问题。如果没有这些困难，则管制者可以通过影响 a_1^F 和 a_2^F 来最大化自己的期望效用（只需要满足公司的个人理性约束）。该结果就是大家所知道的最优结果。在完全信息和“强制性”合同下可以达到这个结果。管制者会具体规定它所想要公司采取的行动，并确保当且仅当公司采取这些行动时，公司会获得保留效用。否则，公司就会被惩罚，因此公司会被迫按照被期望的行事。

对私人信息和不可观察行动两方面的考虑反映了两类问题。当公司的行动不可观察时，并且当其行为对其效用有较大影响时，就会存在道德风险问题。例如，在拉丰和梯若尔（Laffont and Tirole，1986）关于国有企业的模型中，政府不能观察到管理者为了降低成本所付出的个人的高成本的努力。这里也存在道德风险的问题，因为政府要诱使管理者提供“适当”水平的努力，成本是很高的。

当公司了解关于其能力（例如，其技术）的重要信息，而该信息不为管制者所知时，就存在逆向选择的问题。例如，在巴隆和梅耶森（Baron and Myerson，1982）的模型中，被管制的公司完全知道其成本结构，而管制者却不知。

10. 什么时候最优结果是可行的？

即使潜在地存在道德风险和逆向选择的问题，管制者也有可能能够实现最优结果，注意到这点是很重要的。如果满足以下条件，则理想结果可行：（1）公司对货币收入是风险中性的；（2）管制者和公司一开始时拥有相同的信息（即当 R 被宣布时，$I^R = I^F$）；（3）$\tilde{P}$ 不受限制；（4）管制的结果（X）是可观察的。在这种情况下，通过将公司的补偿机制等价于管制者的效用函数（即 $P = U^R$（·）），就可以得到最优结果。这样，公司的期望报酬就等于管制者的期望效用水平。此外，为了实现所希望的租的分配，在生产之前，公司会被要求向管制者一次性支付一笔款项，该笔款项等价于其在该制度下的期望收入和其保留效用水平之间的差值。由于信念

是对称的，管制者和公司将会在这个一次性支付款项（k）的数额上达成一致。并且由于公司是风险中性的，没有必要为其承担的风险支付风险贴水。

当使公司的目标函数和管制者的一样时，且管制者了解公司的私人信息，则公司显然就会按照管制者要求行事。这样，任何的道德风险的问题和/或逆向选择的问题都可以避免。[24]这正是勒布和马盖特（Loeb and Magat，1979）论文中提出的补贴方案的直观解释。他们注意到如果对公司的补贴等于其创造的消费者剩余水平，则公司会实现边际成本价格，并参与到社会最优水平的成本削减活动中来。

为什么要实现最优结果几乎是不可能的？除了我们在第一部分讨论的很重要的一般均衡效应外，在实际中，还有许多的其他原因。第一，公司（或其投资者）可能在一定程度上是风险厌恶的。第二，在管制制度开始实施时，公司掌握的关于所处环境的信息可能比管制者的信息更好。第三，在知道需求和成本函数之前，公司可能并没有资源支付必需的那笔款项，并且对公司负债的法律限制使得处于不利状况的公司可以拒绝支付承诺的一次性款项（k）。当这类的破产约束有效时（即当（RP）中 $\tilde{P}$ 的定义对补偿机制的约束有效时），管制者会根据社会目标来支付公司（即使 $P=U^{R}(\cdot)$），并在公司利润太低的时候接受一个低于 k 的款项，这将不再是最优的。那样的政策给公司提供了不适当的激励，因为无法强迫公司承担其活动的全部“下侧”风险（“downside” risk），例如，公司可能采取一些在管制者看来风险太高的行动（Stiglitz and Weiss，1981）的文章对这点进行了更详细的讨论）。[25]此外，当继续生产的净社会剩余为正的时候，宣布破产可能对公司也存在吸引力。为了消除这种倾向，管制者将设计一种政策以产生一种不同于最优结果的结果。这种扭曲的性质和目的类

---

24 斯蒂格利茨（1974）在他对佃农问题的分析中就提出了该观点。Harris and Raviv（1979）、Holmstrom（1979），以及 Shavell（1979）的一般性理论框架中也提出了该观点。Riordan（1984）对存在需求不确定性的高峰负荷定价问题的解，也依赖于公司的风险中性和关于管制环境的最初信念的对称性。

25 Stiglitz and Weiss（1981）证明了当存在破产约束时，贷款人最优的策略可能应该是拒绝向潜在的借款人提供贷款，而不是仅仅增加借款的费用（也就是提高利率）。他们的基本观点是利率的增加（只有在不破产的情况下才会被支付）会导致：拥有成功可能性较大的项目的借款人面临的期望借贷成本的有效增加会高于那些更容易宣告破产的借款人所面临的。因此利率的增加可能只会起到吸引更少自己希望合作的借款人的作用。此外更高的利率可能会诱使借款人实施更具风险性的项目。

似于当公司的初始信息要优于管制者时会出现的扭曲的性质和目的。[26]稍后我们再讨论这些扭曲。

当管制制度开始实施，公司掌握的关于环境的信息严格优于管制者的信息时，管制者的问题变得更复杂了。许多最近的文献关注的正是这个问题。在这种情况下，管制者必须使管制合同的条款对公司有足够的吸引力，不管公司拥有怎样的私人信息（从数学分析上看，对所有的 I^F 的值，（*RP*）中的个人理性约束（*R*）都必须被满足）。一般而言，在有利的情况下（例如，低成本或高生产率），这将使公司获得一些租。

11. 审核

显然，如果管制者可以通过一些高成本的调查（例如，审核公司的成本）来直接核实公司的私人信息，这样，公司的租通常会减少。例如，众所周知的，如果发现公司提供的关于其私人信息的报告（a_1^F）不真实，就可以给公司一个没有上限的处罚，那么，为了确保公司总能如实地提供自己的私人信息，只需要以任意低的概率实施对公司报告的完全审核（a_1^R）。[27]这样，除了为审核公司的私人信息所需要支付的成本以外，管制者能够实现最优结果。

当然，这个结论可能对审核的质量非常敏感。如果公司获得的补偿取决于对其私人信息并不完全的审核，则公司必然要承担一定的风险。如果公司是风险规避的，则不能实现最优结果。如果公司是风险中性的，非完全信号也可能使管制者实现最优结果（Riordan and Sappington，1985b）给出了实现这种情况的必要条件和充分条件。[28]有一个充分条件（虽然不是必要条件）是信号空间的维数至少要和公司的私人信息空间的维数一样多。

---

26 Sappington（1983a）详细地描述了这些扭曲，而 Kahn and Scheinkman（1985）大体介绍了这些扭曲的特点。

27 不仅在管制现象中，而且在其他许多现象中都会观察到这点。例如，Becker（1968）注意到以很小的概率对犯罪行为实施足够大的惩罚就能够阻止犯罪行为。Stiglitz（1975）证明了如果能从代理人的表现中完全推测出其行动的概率为正（即使任意的小），则可以无成本地解决道德风险问题。Townsend（1979）也记录了对其一般性的观察。Riordan and Sappington（1985b）证明了即使管制者审查的结果不是公开的，导致公司的支付只能取决于管制者宣称的审查发现，而不是直接取决于发现本身，基本的结论仍是成立的。

28 他们的分析考虑了两种情况：信号是公开的和只有委托人（管制者）能够观察到信号。在后一种情况下，最优结果实现的可能性更小，因为任一要利用该信号的（管制）合同都必须使管理者没有动机不如实描述一个人所观察到的信号。

当审核结果不好时，对公司所需承受的处罚的限制（类似破产约束）也会限制管制者实现最优结果的能力。巴隆和贝桑科（Baron and Besanko, 1984b）讨论了当管制者可以对公司的成本实施一个不完全的审核时，以及当对公司的支付存在下限时，管制者的最优策略怎样。直观地讲，为了最大化期望剩余减去审核成本的值，管制者的最优选择是当公司报告的成本高于某个临界值时，就审核其报告，并且当审核结果和初始报告的结论非常不一致时，便对公司施加可能的最高罚款。

当道德风险的问题使得激励问题变得更复杂时，对不完全成本信息的最优使用就有一点不同了。拉丰和梯若尔（Laffont and Tirole，1986）的文章讨论了公司管理者可以通过付出努力来降低经营成本的情况。虽然管理者所付出的努力的程度是不可观察到的，但努力会带来负效用是众所周知的。管理者知道如下私人信息如果不出努力降低成本，生产的边际成本会是多少。因此该模型既包括了逆向选择的问题，也包括了道德风险的问题。在开始生产以后，管制者观察到有关实际生产成本的信息。

为了诱使管理者努力降低成本，管制者必须向管理者承诺：如果实际成本比预期成本低，就会奖励管理者。另一方面，管理者预测的成本比较低时也会获得奖励，否则，就存在夸大预测成本的激励。拉丰和梯若尔证明了，倘若所需的激励是要使管理者如实地报告预测成本并努力降低成本，则最优激励机制是预测成本和实际成本的差值的线性函数。预测成本越小，则奖金增加得越快。

12. 自选择

如果不能使用审核的方法，管制者通常会使用其他一些政策工具。许多研究分析过如何最优地使用价格、补贴和税收来处理逆向选择的问题。大部分研究都假设只有一个技术参数（例如，生产的边际成本），该参数为公司所知，而管制者只知道该参数的统计分布（例如，Baron and Myerson, 1982）。[29]假设管制者和公司之间对该参数的交流是没有成本的。根据（*RP*），这些模型用 a_1^F 表示公司向管制者提交的关于技术参数的报告，而 a_1^R 可能是管制者所做出的关于公司的经营是否应该被终止的决定。如果公司生产，a_2^F 表示公司按照管制价格进行的生产，而 a_2^R 表示对公司产出的确认和对公司的补偿支付。因此公司获得的补偿取决于公司最初向管制者提交的报

29　相关文献中的许多模型都用了相同的假设。例如，参见 Rothschild and Stiglitz（1976）以及 Stiglitz（1977）讨论的保险市场的问题，Stiglitz（1974）所讨论的分成制问题，Mirrless（1975）所讨论的最优税收问题，以及 Stiglitz and Weiss（1985）对信号模型和甄别模型所做的一般性的讨论。有两个或更多参数的自选择模型在数学中会复杂很多。

告和其他一些可观察的变量（例如，生产水平）。[30]这类模型中的绝大多数（除了前面提到的一些例外）都假设无论管制者付出多大的成本，公司的成本都是不能被观察的。

这类模型的解有许多共同的、重要的特点。第一，公司通常都能从它的信息优势中获得租；第二，通常会导致生产决策的扭曲；第三，通常拥有信息（例如成本）不同的公司的行为也不同（例如，产出水平不同）。这些特点源于一些基本的考虑。简单而言，这里假设的唯一信息不对称是两个可能值（$c_1 < c_2$）中的哪一个值代表了公司的（常数）平均生产成本。同样它也假设管制者的目标是最大化消费者剩余的期望值。

在理想的情况下，管制者会让成本是 c_i 的公司以价格 $p_i = c_i$，$i = 1, 2$ 生产出 $Q(c_i)$ 的产品，此处的 $Q(\cdot)$ 是需求曲线。然而，如果有两种价格可以选择，低成本的公司会选择 p_2，并因此而获得租。为了诱使低成本的公司选择低价格，需要向该公司承诺额外支付给它至少等于这些租的款项。管制者为了降低向低成本公司提供的租，最优的办法是抬高 p_2，使之高于 c_2，并因此降低产量，减少该公司利用成本优势的机会。[31]在极端的情况下，最优水平的 p_2 可能非常高，以至于高成本的公司完全不生产。但是管制者对公司的成本为 c_2 越确认，则低成本公司的成本优势越小（也就是 $[c_2 - c_1]$ 的值越小），最优价格水平 p_2 就会越接近边际成本 c_2。[32]

因此，最优管制政策会诱使公司设定越来越接近成本的价格，这样，

30 一些文献中的研究没有将公司和管制者之间一开始的交流正式模型化。例如，对公司的支付只取决于管制的最后结果。如果只有一家公司，且在一开始就掌握了完全信息，并且如果管制者在公司做出生产决策以前不采取任何行动，则这样分析也不会失去一般性。然而，如果缺少这些条件，一开始的无成本的交流会增加管制者的期望收益。例如，在 Nalebuff and Stiglitz（1983a），Weitzman（1976），Holmstrom（1982a）以及 Demski et al.（1985）的模型中，这点都成立，稍后将对这些模型做更详细的介绍。

31 从数学分析的角度而言，因为管制者的目标函数在最优结果得到最大值，对最优价格（或产出水平）的很小偏离，对管制者的效用水平没有影响（根据包络定理可得）。因此，如果除了最低成本（这里是 c_1），对每个实际成本水平，公司都有更低的成本水平，那么使价格稍微高于最优（边际成本）价格，会是限制公司的租的最好办法。

32 例如 Baron and Myerson（1982）和 Sappington（1982，1983a，b）都详细而综合地讨论了管制情形下这种基本的直觉判断。最优税收文献（例如，Mirless 1971；Atkinson and Stiglitz 1976）、保险文献（例如，Rothschild and Stiglitz，1976；Stiglitz，1977），以及对非线性定价的研究（例如，Goldman，Leland and Sibley，1984；Maskin and Riley，1984）中的相应结论也包含了这点。Copper（1984）描述了具有这种性质的逆向选择问题的解的一般特征。

之后，管制者就可以通过公司设定的价格来推断其真实成本。然而，只有当管制者能够做出可信的承诺，承诺其不会利用它最后推断出来的成本信息来对付公司时，这样的“自选择”才会出现。在这部分结束的时候，我们会再讨论当管制者缺乏这种事前承诺的能力时，最优管制政策会如何变化。

当管制者的政策工具更多时，除了高于边际成本以外，通常还会出现其他价格扭曲。例如，如果由于公司的技术参数的不同，可变生产成本受可观察固定成本的影响不同，则管制者通常会命令公司使用非成本最小化的固定成本进行生产（Sappington，1983b）。[33]更一般情况的是管制者会根据可观察的投入和/或产出来决定公司最后得到的补偿，而投入和/或产出在很大程度上是同公司的私人信息有关的。这样做通常会使投入的使用以及产出发生扭曲，偏离最优水平。并且即使观察这些间接感兴趣的变量并不是完全准确的，管制者还需要支付一定的成本，但只要能获得足够多的、有关其关心变量的信息，管制者就可能会支付这样的成本。[34]

13. 阿弗奇—约翰逊

已在管制文献中受到广泛关注的一种投入扭曲是阿弗奇和约翰逊（Averch-Johnson，A-J，1962）讨论的过度投资。他们的分析是如何与（*RP*）中给出的一般性模型联系起来，这是一个很有意思的问题。在 A－J 模型中似乎有两个关键的信息不对称。其一，公司知道投入价格（也就是资本的成本 r）但管制者不知道。其二，对于其技术的函数形式，公司拥有信息优势。我们可能觉得管制者会认为 r 在某区间（$\underline{r}$，$\bar{r}$）上有正密度。但事实上，管制者总是设置高于资本真实成本的容许回报率 s，这说明管制者设置 $s=\bar{r}$。如果强迫公司破产的成本是无穷大的，则该政策可能是最优的。巴隆和塔戈特（Baron and Taggart，1980）也证明了 A－J 管制政策只有当管

33 Besanko（1984）还提出了从技术扭曲中得到的收益。

注意，不管公司的私人信息在一开始的时候是否是完全的，都会有这些扭曲。只要公司的初始信息优于管制者的，最优策略就是诱导可观察的决策变量产生扭曲。具体的分析参见 Sappington（1982，1984）。

要注意到，还有一个原因也导致了生产技术被故意扭曲。以污染控制为例，公司可能会命令公司应用某种技术（例如，洗涤器），而不对污染征税。虽然这种直接控制是没有效率的，因为其使得公司不能利用自己在技术方面的信息优势来实现最优的污染消除量，当监控污染排放量的成本太高，以至于间接控制不可行时，直接控制就会优于间接控制。

34 该类信息的价值似乎存在一种非凹性（Radner and Stiglitz，1984；Singh，1985）。所以管制者不会“购买”和自己感兴趣的变量关联度不大的信息。

制者不知道公司的成本结构时才是最优的。否则，最优的管制政策应该将公司对股本总额的选择看作是可以传递其技术信息的信号。[35]

A－J模型中的一个关键假设是，r是外生的。事实上公司的资本成本可能对管制政策是敏感的。一种可能的情况是，如果s低于当前的资本成本，公司的规模会降低（因为筹集新资金的困难），因此会降低公司的经营风险，从而降低公司的资本成本。第二种可能的情况是公司提高其绩效（例如，降低组织的“松懈度”），这将降低公司的资本成本。所以潜在假设——如果s低于r，公司将会破产——可能是比较极端的情况。

实际中，管制者会将对回报率的管制和对公司的资本选择的直接控制这两种手段结合起来使用。霍姆斯特姆（1984）的分析说明了当管制者和公司之间的目标的差异越小时，以及当双方的信息不对称越严重时，则这些控制就越不严厉。[36]当公司的信息比管制者更充分时，则公司更有能力决定社会最优的投资水平。因此，倘若其利益和社会利益之间的差异不是非常大，在投资选择方面，公司将会被授予更大的自主权。

14. 重复性相互作用

至此，在我们已关注的模型中，管制者和公司之间的相互作用都是非重复性。实际中，管制常常处在不断变化中。因此我们现在转而讨论重复性的相互作用是如何改变静态模型推导出的结论的。其中的核心问题是：如果当未来的管制取决于现在的表现，公司会如何改变当前的行为来影响未来的管制。因此这里的主要问题类似在计划经济中和设定劳工报酬时遇到的问题。在前苏联式激励问题中（例如，Holmstrom，1982a；Loeb and Magat，1978；Weitzman，1976），当未来的产量目标随着当前产量的增加而增加时，公司就可能故意降低自己的产量，而不是充分利用自己的生产能力。类似地，如果未来的计件工资率会因为现在生产速度的增加而被下调，则工人会降低现在的努力程度（Stiglitz，1975）。

首先，我们注意到在一些条件下静态模型的结论同动态模型的结论是一样的。例如，当管制者可以无成本地约束自己严格执行在管制制度开始

35 Besanko（1984）证明了如果A－J模型中的关键的技术信息不对称可以被减少到只对一个参数不对称（例如，Baron and Myerson，1982），那么最优管制政策会让允许的回报率随着公司受到投入资本的量反向变化，由此可以调节公司过度投资的倾向。

36 Holmstrom（1984）对委托问题（单个代理人时）的分析可以被视做是（RP）的一个特例。规定代理人的可观察行动（例如，他对资本量的选择）必须在一定范围内，等同于假设如果管制者观察到公司的行动在该范围外，就会对该公司实施无穷大的罚款。

时宣布的管制政策条款，则此时的逆向选择问题同静态模型中一样。此外，当公司的成本结构（它自己完全了解）在每一期都一样，而管制者对这些成本的不完全信息在每一期也一样时，管制者无法观察到公司的历史成本，则动态和静态的情况一样。在这些情况下最优的动态管制政策将是在管制的每一期都制定单独的最优政策。[37]当公司各期的成本并不是完全相关时，最优动态管制政策将不同于静态模型的，虽然两者中的产出扭曲的性质一样。[38]

如果管制者能够观察到公司过去的表现，则管制者和公司之间的成本信息的严重不对称或许不会导致重复性逆向选择问题。夏平顿和希伯莱（Sappington and Sibley，1985）设计了一种管制机制，能够保证总剩余最大化，并且公司在每一期都获得它的保留利润。管制者只需知道公司的需求曲线和贴现率，并且能够滞后一期（例如一年）观察到公司的历史支出（公司可能会策略性地使支出超过最小可能生产成本），管制者不需要知道公司的技术。该机制在每一期把公司活动产生的消费者剩余的增加奖励给公司，并对公司征收一个等于其前一期的收入和支出之差的税收。因为它使得管制者和公司的目标在本质上是一致的，所以它是有效的。[39]

观察历史支出的能力最后甚至可以导致不合逻辑的结果——管制者和公司之间的信息不对称更严重。福格桑和凡辛格（Vogelsang and Finsinger，1979）讨论了管制者对一家生产多种产品的垄断公司的技术或需求结构完全不了解的情况。[40]他们提出了一个非常简单漂亮的机制，使得公司的定价收敛于拉姆齐定价。该机制限制公司在每一期的价格的拉斯贝尔指数（Laspeyre's index）都不能超过前一期的实际支出。该指标的权数是公司在

37 Rothschild and Stiglitz（1976）也提到了这点，而 Baron and Besanko 也具体分析了这点。

38 例如，Baron and Besanko（1984a）分析了 Baron and Myerson（1982）的模型扩展到了多期的情况，并且未来的成本和现在的成本有正相关关系，虽然并不是完全相关。他们发现最优管制政策仍存在扭曲，以便限制公司的租，不过通常在第一期以后，扭曲的程度就开始减弱了。

39 该机制的逻辑类似于 Loeb and Magat（1979）的基本观点。虽然管制者观察历史支出的能力解决了在消费者和公司之间如何分配租的问题，但我们应该注意到，如果公司的所有权和管理权没有分离，且如果超过最小生产成本的支出能够给所有者—管理者提供直接效用，那么 Sappington and Sibley（1985）提出的机制对该问题的解决是不完全的。

40 从数学分析的角度，唯一需要知道的是公司的成本结构是递减的射线型平均成本（Panzar and Willig，1977）。（这个假设确保了在被提议的管制机制下，公司的利润总是非负的。）尤其要注意的是（Loeb and Magat，1979；Sappington and Sibley，1985 的模型所提出的），关键的信息不对称不能被减少到只是一个技术参数的不对称，且公司知道该参数的实际值，管制者知道其分布。所以这些模型中的信息不对称问题更严重。

前一期生产的产品数量。因此公司在任一期的实现利润越大，后来的价格指数受到的限制就越大。最后，消除了利润，而最大化了总剩余。[41]

同样地，重复性相互作用和过去表现的可观察性也可以解决道德风险问题。雷德勒（Radner，1984）以及林哈特、雷德勒和辛德（Linhart，Radner and Sinden，1983）证明了倘若管理者的贴现率足够低，并且他被解雇的个人成本足够大，就可以在不给其施加不适当风险的情况下，促使他为管制者的利益而行动。[42]一般对其结论的直观的解释是：在重复博弈下，大数定理使管制者可以更准确地区分公司自身努力和勤劳的缺乏与“坏运气”（其在长期内会在统计上被“好运气”所抵消）这两种情况。因此管制者更能促使公司按它所希望的行事。例如，管制者可以威胁公司：如果它认为公司长期的表现太差，就要终止公司现有管理工作。

15. 代理人之间的比较

重复性相互作用可以形成多个观察，使得管制者能够更好地控制公司的行为。另外一种形成多个观察的方法是比较两个或多个代理人的表现。对管理者而言，相对表现是一个非常有用的工具。例如，德穆斯基和夏平顿（Demski and Sappington，1984）的研究表明，当管制者监管的是两个风险中性的、成本相互关联（然而只有些微的关联）的公司时，可以得到最优结果——即使在一开始的时候公司完全知道其成本信息，且没有和管制者分享该信息。在一与之相关模型中，纳尔巴夫和斯蒂格利茨（Nalebuff and Stiglitz，1983a）证明了在被管制的、拥有私人信息的公司之间，最优结果是如何作为纳什均衡被实现的，即使这些公司都是风险规避的。

根据每家公司的相对表现，对每个公司给予奖励的管制机制可以减轻使公司消极行事（例如，夸大成本或者怠工）的激励。实际上，一家公司的表现是另一家公司行为的一种自然监督。当然，这种监督的价值常常取决于这些公司的特点（例如，技术和需求函数）有多大的相似度。相

41 就如Sappington（1980）所证明的那样，该机制对公司的策略性行为（例如，故意成本膨胀）也不是免疫的。但是可以证明在这类行为下，作者的基本结论还是成立的。尤其是，价格还是会收敛到拉姆齐水平，虽然收敛速度可能会慢一些。Finsinger and Vogelsang（1981，1985）也提出了引人注目的机制，促使国有企业的管理者利用成本信息和需求信息优势来形成最优定价决策。上述机制不要求使用成本信息或需求信息，而是每期都把近似地等价于管理者创造的总剩余的增加值奖励给管理者。Tam（1985）对这种类型的机制做了一个很有意思的综述。

42 对该结论的准确陈述以及直观解释参见Radner（1981）的分析。

似度越高，从相对表现中推导出来的无法观察的公司行为就越多。由于相对表现可以帮助减轻信息不对称所导致的问题，管制者甚至可能发现允许新的公司进入存在自然垄断的行业也是有好处的。德穆斯基等人（Demski et. al.，1985）的论文中也讨论了纳尔巴夫和斯蒂格利茨（Nalebuff and Stiglitz，1983a）提出的这种可能性存在。他们证明了最优管制政策可能是不时地用成本更高的新入者来代替现任的公司，以减弱在位者夸大成本的倾向。

我们还注意到，事前有多家潜在生产者为成为被管制的垄断者而竞争，可以限制事后垄断公司的租（Demsetz，1968；McAfee and McMillan，1984）。[43]例如，勒布和马盖特（Loeb and Magat，1979）讨论过如果根据垄断公司产生的消费者剩余来对其进行补贴，并且如果存在足够的竞争来争取该垄断权，则对垄断权的竞标就能实现最优结果。里奥顿和夏平顿（Riordan and Sappington，1985a）证明了当潜在的生产者相互独立地预测垄断特许权的期望价值时，"足够的竞争"会导致有无数的竞标者。潜在生产者会少一些给赢标者带来租。此外，为了限制这些租，通常管制者的最优策略是诱导生产者将产品的价格定得严格高于边际生产成本。生产上的扭曲使得竞标的竞争更激烈，因此可以限制垄断者的租（相关的讨论，参见 Laffont and Tirole，1985；McAfee and McMillan，1986）。如果对垄断特权价值的最高私人评估越低（也就是"中标价"越小），则通常扭曲的程度会越大。

下面是对这些结论的一个简单的解释。当实际生产成本很小时，生产上的扭曲限定了生产者能够获得的租。扭曲越显著，则租的减少也越大。因此，通过承诺实现越严重的扭曲，则中标价越低，管制者就会劝潜在的生产者在报出他们对垄断特权估值时不要保守。低报其真实的估值，不仅使公司承担了更大的可能将特权输给竞争对手的风险，而且在赢得特权后，它的标价越低，它会采取获利性越小的经营方式。

注意到对垄断特权的竞标使管制者能够处理两个难题。标价既可以限制最后的垄断者的租，又能有助于识别出最适合承担该生产任务的公司。直观而言，成本最低的公司对垄断特权的出价会最高。[44]当然更适合的成本会促

43 Williamson（1975）引用了各种理由来说明为什么很难进行成功的特权拍卖。

44 最优的特权拍卖不一定保证生产成本最低的公司会得到特权。例如，如果公司采取混合竞标策略，中标的可能并非是成本最低的公司。而且，如同梯若尔所指出的，当最优激励机制包含"混合"或"捆绑"解时，特权被授予给非成本最低生产者的概率为正（也可以参见 Myerson，1981）。

进更高的出价。然后，参与者更大的相似性通常会使得管制者可以更加近似地实现最优结果。

一般而言，具体给出存在多个公司时的最优管制政策是件困难而复杂的事情。例如，实施让公司进行简单的比赛，奖励表现最好的公司的管制政策，当公司之间的差异（如能力）相对于公共环境冲击（例如关键要素价格的变化）变得更小，这种管制政策就没那么有效率了。[45]纳尔巴夫和斯蒂格利茨（1983b）已证明了对"失败者"实施惩罚要比奖励"成功者"更具激励效应。他们还证明了在更复杂的激励机制中，可能会存在多个均衡，也可能一个均衡也不存在。关键是要避免公司过多地进行模仿。如果公司因为失败会得到很严厉的惩罚，而由于表现好而获得的奖励很少，那么对公司有效率地承担风险的激励就会减少，技术发展也因此而受到阻碍。这个和其他的相关问题应该得到更多的关注。[46]

16. *有限承诺*

在结束前，我们来简单讨论一个此前都一致被回避的重要问题。这个问题是：当管制者无法事前承诺它在某一时期一定会遵照已宣布的政策行事时，管制政策会有怎样的必要变化。回顾一下我们此前讨论过的，如果管制者可以做出可信的承诺——不会利用从公司当前表现推导出来的信息来采取对公司不利的行动，则管制者就能够激励公司在当前表现得更好。例如，当有这样的承诺时，公司的成本越低，它就会被诱导在每一期生产更多的产品，因此会增加消费者剩余。然而，当公司这样做不会获得很大的补贴，而且管制者之后会降低它在未来能够获得的租时，公司不会让自己被认出是低成本生产者。因此，当管制者事前承诺的能力减小时，使公司的绩效和其能力相称（也就是达到"分离"状态）所需付出的成本也会增加。

当管制期限是无限期的，并且公司的贴现率为零的极端情况下（例如，Roberts，1984），不会或者实际上是不能形成分离的。在这些条件下要达到分离，只能给公司提供一个无限大的贿赂。当管制期为有限期并且/或者是正的贴现率时（Britto et al.，1985；Baron and Besanko，1985），达到分离结果所要求的支付不会是无限大的。因此最优管制政策会使初始的一段时间没

---

45 Lazear and Rosen（1981）和 Green and Stokey（1983）都得出了该结论以及相关的结论。

46 Holmstrom（1982b）和 Mookherjee（1984）的分析针对为多个代理人设计激励机制的问题，提出了一些很好的观点。

有分离结果，在之后会实现分离。[47]

因此，当可以无成本地做出承诺时，用持续的扭曲（例如价格超过边际成本）来诱使分离情况的出现是最优的政策，当没有能力做出承诺时，最优管制政策会在消除所有扭曲之前的一段时期内导致更严重的扭曲。[48]

17. 对未来研究的一些建议

至此，我们已经回顾了文献中对信息不对称情况下设计管制政策的讨论。在结束本文之前，我们来关注一下一些值得进一步研究的问题和现象。

首先，应注意到我们的分析（同很多文献一样）集中于一种特殊的委托—代理关系。然而，还有其他一些有趣的、类似的关系也构成了管制情形的特征。例如，被管制公司的所有者必须为其管理者构造合适的激励，使其为所有者的利益行事。类似地，管理者要激励工人采取恰当的行动。此外，国会可能有必要为管制者提供足够的激励，使管制者努力地完成工作。在我们完全理解这么多的委托—代理问题之间的微妙的相互关系之前，还有大量的研究需要完成。[49]

相关的研究关注如何控制同时被几个管制者所管制的公司。当管制者的利益不完全一致时，就会产生管制命令之间相互冲突的问题。管制权重叠时的基本原理以及这种重叠下的最优激励机制的构建都值得进一步讨论。[50]出于政治方面的考虑是出现这种管理权分散的原因之一。另一个原因可能是设

47 Baron and Besanko（1985）的模型指出，如果在任何时候不存在个人理性约束条件，则初始阶段或者后面的分离阶段的持续时间会缩减到零。

Freixas et. al.（1985）也分析过一个有趣的两期模型，在该模型中，管制者缺乏承诺能力。该分析清楚地指出了当管制者不能对将来的激励机制做出承诺时，会导致的损失。另一个相关的很重要的多期模型是 Lewis（1986）提出的模型。

48 在管制情形下，需要考虑是否有一些形式的承诺要比另一些更容易实现。例如，Sappington（1986）建议当承诺使用一种监督技术（也就是观察公司实际成本的花费）比承诺任一已有技术的使用频率更容易时，承诺使用无效率的监督技术可能会带来收益。这属于用管制层级中某阶段的无效率，来补偿另一阶段承诺能力缺乏的情况之一。

49 关于这方面的研究，请参见 Demski and Sappington（1986）。

50 Baron（1985）的研究在这方面作出了很大的贡献。他分析了当公共事业单位机构和环境保护机构之间的目标出现冲突时的最优管制政策的设计。

计合适的管制政策所需要的信息使得最好存在几个不同的管制者。[51]

虽然我们回顾的文献观察到了有效率的管制机制的一般性质，但通常都没有给出具体的政策建议。例如，虽然我们知道公共电力公司的管理者一般应该被迫承担一些社会负担——这些负担是由于有关核能厂的决策失误造成的，但其具体应该承担的负担是多少却不是很清楚。如何在纳税人、股东、管制者以及公用事业单位的管理者之间最优地分配风险，还值得更仔细地研究。如前所述，由于存在大量的信息不对称以及委托—代理问题之间的相互关联，使得这个问题非常复杂。

另一方面的信息不对称也限制了对关键的管制问题的具体政策制定。因为大多数的经济学家缺少至关重要的、有关政策实施的特定行业的制度信息，他们不能提出准确的（或者至少会轻率地提出）政策建议。所以，很重要的一点是大家能够分享自己所掌握的有关理论分析和制度细节方面的信息。

参考文献

Arnott, R., and Stiglitz, J. 1985. "The Welfare Economics of Moral Hazard." Princeton University mimeo. March.

Atkinson, A., and Stiglitz, J. 1980. *Lectures in Public Economics*. New York: McGraw-Hill.

Atkinson, A., and Stiglitz, J. 1976. "The Design of Tax Structures: Direct vs. Indirect Taxation." *Journal of Public Economics* 6: 55 – 75.

Atkinson, A., and Stiglitz, J. 1972. "The Structure of Indirect Taxation and Economic Efficiency." *Journal of Public Economics* 1: 97 – 119.

Averch, H., and Johnson, L. 1962. "Behavior of the Firm under Regulatory Constraint." *American Economic Review* 52: 1052 – 1069.

Baron, D. 1985. "Noncooperative Regulation of a Nonlocalized Externality." Stanford University Discussion Paper. February.

Baron, D. 1984. "Regulatory Strategies under Asymmetric Information." In M. Boyer and R. Kihlstrom, eds., *Bayesian Models in Economic Theory*. Amsterdam: North Holland, pp. 155 – 180.

Baron, D., and Besanko, D. 1985. "Commitment in Multiperiod Information Models." Working Paper at Stanford University. May.

Baron, D., and Besanko, D. 1984a. "Regulation and Information in a Continuing Relationship." *Information, Economics and Policy* 1: 267 – 302.

---

51 我们的综述局限于当公司相对管制者拥有信息优势的情况，但也有可能管制者在管制环境的某些方面相对公司也拥有信息优势。Myerson（1983）提出了当委托人信息更多时的最优激励机制的设计。Braverman and Stiglitz（1982）和 Bernheim and Whinston（1984）也提出了当多个委托人对一个代理人时的最优激励机制设计。

Baron, D., and Besanko, D. 1984b. "Regulation, Asymmetric Information and Auditing." *Rand Journal of Economics* 14: 447 – 470.

Baron, D., and Myerson, R. 1982. "Regulating a Monopolist with Unknown Costs." *Econometrica* 50: 911 – 930.

Baron, D., and Taggart, R. 1980. "Regulatory Pricing Procedures and Economic Incentives." In M. Crew, ed., *Issues in Public Utility Economics and Regulation*. Lexington, Mass.: Lexington Books.

Baumol, W., and Bradford, D. 1970. "Optimal Departures from Marginal Cost Pricing." *American Economic Review* 60: 265 – 283.

Baumol, W., and Klevorick, A. 1970. "Input Choices and Rate-of-Return Regulation: An Overview of the Discussion." *Bell Journal of Economics* 1: 162 – 190.

Baumol, W., Panzar, J., and Willig, R. 1982. *Contestable Markets and the Theory of Industry Structure*. New York: Harcourt, Brace, Jovanovich.

Beeker, G. 1968. "Crime and Punishment: An Economic Approach." *Journal of Political Economy* 76: 169 – 217.

Berle, A. and Means, G. 1932. *The Modern Corporation and Private Property*. New York: Macmillan.

Bernheim, D., and Whinston, M. 1984. "Common Agency." Harvard Institute of Economic Research Paper No. 1074.

Besanko, D. 1984. "On the Use of Revenue Requirements Regulation under Imperfect Information." in M. Crew, ed., *Analyzing the Impact of Regulatory Change*. Lexington, Mass.: Lexington Books, pp. 39 – 58.

Boiteaux, M. 1956. "Sur la Question des Monopoles Publics Astreints a l'Equilibre Budgetaire" *Econometrica* 24: 22 – 40.

Braverman, A., and Stiglitz. J. 1982. "Sharecropping and the Interlinking of Agrarian Markets." *American Economic Review* 72: 695 – 715.

Britto. R., Hamilton, J., Slutsky, S., and Stiglitz, J. 1985. "Taxation and commitment." Mimeo. June.

Cooper, R. 1984. "On Allocative Distortions in Problems of Self-Selection." *Rand Journal of Economics* 15: 568 – 577.

Dasgupta, P., and Stiglitz, J. 1985. "Sunk Costs, Competition and Welfare." Princeton University Working Paper. July.

Demsetz, H. 1968. "Why Regulate Utilities?" *Journal of Law and Economics* 7: 55 – 65.

Demski, J., and Sappington, D. 1986. "Hierarchical Regulatory Control." Bell Communications Research mimeo. July.

Demski, J., and Sappington, D. 1984. "Optimal Incentive Schemes with Multiple Agents." *Journal of Economic Theory* 33: 152 – 171.

Demski, J., Sappington, D., and Spiller, P. 1985. "The Disciplinary Role of Switching Suppliers: A Regulatory Interpretation." Working Paper at Bell Communications Research. June.

Ebrill, L., and Slutsky, S. 1985. "Pricing Rules for Intermediate and Final Good Regulated Industries." University of Florida Working Paper, March.

Ekern, S., and Wilson, R. 1974. "On the Theory of the Firm in an Economy with Incomplete

Markets." *Bell Journal of Economics and Management Science* 5: 1971 – 180.

Farrell, J. 1985. "Allocating and Abrogating Rights: How Should Conflicts be Resolved under Incomplete Information." GTE Laboratories mimeo.

Finsinger, J., and Vogelsang, I. 1981. "Alternative Institutional Frameworks for Price Incentive Mechanisms." *Kyklos* 34: 388 – 404.

Finsinger, J., and Vogelsang, I. 1985. "Strategic Management Behavior under Reward Structures in a Planned Economy." *Quarterly Journal of Economics* 100: 263 – 270.

Friexas, X., Guesnerie, R., and Tirole, J. 1985. "Planning under Incomplete Information and the Ratchet Effect." *Review of Economic Studies* 52: 173 – 191.

Goldman, M., Leland, H., and Sibley, D. 1984. "Optimal Nonuniform Prices." *Review of Economic Studies* 51: 305 – 319.

Greenwald, B., and Stiglitz, J. 1986. "Externalities in Economies with Imperfect Information and Incomplete Markets." Princeton University Working Paper. January.

Green, J., and Stokey, N. 1983. "A Comparison of Tournaments and Contracts." *Journal of Political Economy* 91: 349 – 364.

Grossman, S., and Hart, O. 1980. "Takeover Bids, the Free Rider Problem, and the Theory of the Corporation." *Bell Journal of Economics* 11: 42 – 65.

Grossman, S., and Hart, O. 1981. "The Allocational Role of Takeover Bids in Situations of Asymmetric Information." *Journal of Finance* 36: 253 – 270.

Grossman, S., and Stiglitz, J. 1980. "Stockholder Unanimity in Making Production and Financial Decisions." *Quarterly Journal of Economics* 94: 543 – 566.

Harris, M., and Raviv, A. 1979. "Optimal Incentive Contracts with Imperfect Information." *Journal of Economis Theory* 20: 231 – 259.

Holmstrom, B. 1982a. "The Design of Incentive Schemes and the New Soviet Incentive Model." *European Economic Review* 17: 127 – 148.

Holmstrom, B. 1979. "Moral Hazard and Observability." *Bell Journal of Economics* 10: 74 – 91.

Holmstrom, B. 1982b. "Moral Hazard in Teams", *Bell Journal of Economics* 13: 324 – 340.

Holmstrom, B. 1984. "On the Theory of Delegation." In M. Boyer and R. Kihlstrom, eds., *Bayesian Models in Economic Theory*. Amsterdam: North Holland, pp. 115 – 142.

Kahn, C., and Scheinkman, J. 1985. "Optimal Employment Contracts with Bankruptcy Constraints." *Journal of Economic Theory* 35: 343 – 365.

Laffont, J., and Tirole J. 1985. "Auctioning Incentive Contracts." MIT Working Paper. October.

Laffont, J., and Tirole, J. 1986. "Using Cost Observation to Regulate Firms." *Journal of Political Economy*, forthcoming.

Lazear, E. and Rosen, S. 1981. "Rank-Order Tournaments as Optimum Labor Contracts." *Journal of Political Economy* 89: 841 – 864.

Lewis, T. 1985. "Reputation and Contractual Performance in Long Term Projects." Working Paper at the University of British Columbia. June.

Linhart, P., Radner, R., and Sinden, F. 1983. "A Sequential Principal-Agent Approach to Regulation." Bell Laboratories Discussion Paper.

Loeb, M., and Magat W. 1979. "A Decentralized Method for Utility Regulation." *Journal of*

Law and Economics 22: 399 - 404.

Loeb, M., and Magat, W. 1978. "Success Indicators in the Soviet Union: The Problem of Incentives and Efficient Allocations." *American Economic Review* 68: 173 - 181.

March, J., and Simon, H. 1958. *Organizations.* New York: Wiley.

Maskin, E., and Riley, J. 1984. "Monopoly with Incomplete Information." *Rand Journal of Economics* 15: 171 - 196.

McAfee, R., and McMillan, J. 1986. "Competition for Agency Contracts." Working Paper at University of Western Ontario. January.

McAfee, R., and McMillan, J. 1984. "Incentive Contracts and Selection of Agent by Principal." Working Paper at University of Western Ontario. May.

McFarland, H. 1985. "Ramsey Pricing of Inputs with Downstream Monopoly Power and Regulation." EPO Discussion Paper 85 - 2. U. S. Department of Justice.

Mirrlees, J. 1971. "An Exploration in the Theory of Optimum Income Taxation." *Review of Economic Studies* 38: 175 - 208.

Mookherjee, D. 1984. "Optimal Incentive Schemes with Many Agents." *Review of Economic Studies* 51: 433 - 446.

Myerson, R. 1983. "Mechanism Design by an Informed Principal." *Econometrica*, 51: 1767 - 1798.

Nalebuff, B., and Stiglitz, J. 1983a. "Information, Competition and Markets." *American Economic Review* 73: 278 - 283.

Nalebuff, B., and Stiglitz, J. 1983b. "Prices and Incentives: Towards a General Theory of Compensation and Competition." *Bell Journal of Economics* 14: 21 - 43.

Panzar, J., and Willig, R. 1977. "Free Entry and the Sustainability of Natural Monopoly." *Bell Journal of Economics* 8: 1 - 22.

Radner, R. 1981. "Monitoring Cooperative Agreements in a Repeated Principal-Agent Relationship." *Econometrica* 49: 1127 - 1148.

Radner, R. 1974. "A Note on Unanimity of Stockholders' Preferences among Alternative Production Plans." *Bell Journal of Economics* 5: 181 - 186.

Radner, R. 1984. "Repeated Moral Hazard with Low Discount Rates." Bell Laboratories Discussion Paper.

Radner, R., and Stiglitz, J. 1984. "A Nonconcavity in the Value of Information." In M. Boyer and R. Kihlstrom, eds., *Bayesian Models in Economic Theory.* Amsterdam: North Holland, pp. 33 - 52.

Ramsey, F. 1927. "A Contribution to the Theory of Taxation." *Economic Journal* 37: 47 - 61.

Riordan, M. 1984. "On Delegating Price Authority to a Regulated Firm." *Rand Journal of Economics* 15: 108 - 115.

Riordan, M., and Sappington, D. 1985a. "Awarding Monopoly Franchises." Working Paper at Stanford University and Bell Communications Research. Revised December.

Riordan, M., and Sappington, D. 1985b. "Optimal Incentive Contracts with Public and Private Ex Post Information." Working Paper at Stanford University and Bell Communications Research. June.

Roberts, K. 1984. "The Theoretical Limits to Redistribution." *Review of Economic Studies* 51:

177 – 196.

Ross, S. 1973. "The Economic Theory of Agency: The Principal's Problem." *American Economic Review* 63: 134 – 139.

Rothschild, M., and Stiglitz, J. 1976. "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Information." *Quarterly Journal of Economics* 90: 225 – 243.

Sah, R., and Stiglitz, J. 1985. "Human Fallibility and Economic Organization." *American Economic Review* 76: 292 – 297.

Sappington, D. 1986. "Commitment to Regulatory Bureaucracy." Bell Communications Research Working Paper. January.

Sappington, D. 1984. "Incentive Contracting with Asymmetric and Imperfect Precontractual Knowledge." *Journal of Economic Theory* 34: 52 – 70.

Sappington, D. 1983a. "Limited Liability Contracts Between Principal and Agent." *Journal of Economic Theory* 29: 1 – 21.

Sappington, D. 1983b. "Optimal Regulation of a Multiproduct Monopoly with Unknown Technological Capabilities." *Bell Journal of Economics* 14: 453 – 463.

Sappington, D. 1982. "Optimal Regulation of Research and Development under Imperfect Information." *Bell Journal of Economics* 13: 354 – 368.

Sappington, D. 1980. "Strategic Firm Behavior under a Dynamic Regulatory Adjustment Process." *Bell Journal of Economics* 11: 360 – 372.

Sappington, D., and Sibley, D. 1985. "Regulatory Incentive Schemes Using Historic Cost Data." Working Paper at Bell Communications Research. Revised August.

Shapiro, C., and Stiglitz, J. 1984. "Equilibrium Unemployment as a Worker Discipline Device." *American Economic Review* 74: 433 – 444.

Shavell, S. 1979. "Risk Sharing and Incentives in the Principal and Agent Relationship." *Bell Journal of Economics* 10: 55 – 73.

Singh, N. 1985. "Monitoring and Hierarchies: The Marginal Value of Information in a Principal-Agent Model." *Journal of Political Economy* 93: 599 – 609.

Stigler, G. 1972. "The Theory of Economic Regulation." *Bell Journal of Economics* 2: 3 – 21.

Stiglitz, J. 1985. "Credit Markets and the Control of Capital." *Journal of Money, Credit, and Banking* 17: 133 – 152.

Stiglitz, J. 1975. "Incentives, Risk and Information: Notes Towards a Theory of Hierarchy." *Bell Journal of Economics* 6: 552 – 579.

Stiglitz, J. 1977. "Monopoly Non-Linear Princing and Imperfect Information: The Insurance Market." Review of Economic Studies 44: 407 – 430.

Stiglitz, J. 1970. "On the Optimality of the Stock Market Allocation of Investment." Paper presented at the Far Eastern Meetings of the Econometric Society, Tokyo, Japan, June 27 – 29.

Stiglitz, J. 1972. "Some Aspects of the Pure theory of Corporate Finance: Bankruptcies and Takeovers." *Bell Journal of Economics and Management Science* 3: 458 – 482.

Stiglitz, J. 1974. "Risk Sharing and Incentives in Sharecropping." *Review of Economic Studies* 41: 219 – 256.

Stiglitz, J. 1982. "Self-Selection and Pareto Efficient Taxation." *Journal of Public Economics* 17: 213 – 240.

Stiglitz, J., and Weiss, S. 1983. "Alternative Approaches to Analyzing Markets with Asymmetric Information." *American Economic Review* 73: 246 - 249.

Stiglitz, J., and Weiss, A. 1981. "Credit Rationing in Markets with Asymmetric Information." *American Economic Review* 71: 393 - 410.

Stiglitz, J. and Weiss, A. 1983. "Incentive Effects of Terminations: Applications to the Credit and Labor Markets." *American Economic Review* 73: 912 - 927.

Stiglitz, J., and Weiss, A. 1985. "Sorting Out the Differences between Screening and Signaling Models." Bell Communications Research Discussion Paper.

Tam, M. 1985. "Reward Structures in a Planned Economy: Some Further Thoughts." *Quarterly Journal of Economics* 50: 279 - 290.

Townsend, R. 1979. "Optimal Contracts and Competitive Markets with Costly State Verification." *Journal of Economic Theory* 21: 265 - 293.

Vogelsang, I., and Finsinger, J. 1979. "A Regulatory Adjustment Process for Optimal Pricing by Multiproduct Monopoly Firms." *Bell Journal of Economics* 10: 157 - 171.

Weitzman, M. 1976. "The New Soviet Incentive Model." *Bell Journal of Economics* 251 - 257.

Williamson, O. 1976. "Franchise Bidding for Natural Monopolies——In General and with Respect to CATV." *Bell Journal of Economics* 7: 73 - 104.

Williamson, O. 1975. *Markets and Hierarchies: Analysis and Antitrust Implications.* New York: The Free Press.

自由化、银行业道德风险与审慎监管：资本要求管用吗？*

在动态道德风险模型之下，激烈的竞争会改变银行原本审慎的经营行为。尽管资本要求能够约束银行的经营行为并使之更加审慎，但是这种措施所产生的结果是帕累托无效的。资本要求使得银行的资产净值在风险之下，从而降低了银行投机的激励。不过，资本要求同样会有负面效应，因为它会降低银行特许权的价值，从而鼓励了投机行为。将控制银行的存款利率作为附加的监管手段，可以实现帕累托有效的结果，因为这样增加了银行特许权的价值，从而鼓励审慎的经营行为。即便在均衡路径下，银行的存款利率约束不是紧的，它也至少可以在均衡路径之中消除投机行为。（*JEL* *G*2，*E*4，*L*5）

银行危机的爆发有愈演愈烈之势。在过去的20年间，重大的银行危机爆发的频率越来越高。银行危机之所以危害巨大，不仅仅是因为它给经济之中的某个特别部门带来了巨大的灾难，而主要是因为它对整个经济所产生的冲击。在19世纪，美国经济所经历的衰退绝大多数都与金融恐慌有关。同样不能忽视的是，银行危机对政府预算所造成的影响，因为政府通常在解决危机的过程中承担了相当大的成本。世界银行在分析了过去20年之中的许多例子后认为，这种成本可以达到国民生产总值的百分之四十之多。也许最著名的例子就是美国曾经爆发的储蓄信贷协会危机（S&L），总共造成了 大约1 800亿美元的损失，占国民生产总值的3.2%，以及现在日本所经

* Thomas F. Hellmann, Kevin C. Murdock, Joseph E. Stiglitz, "Liberalization, Moral Hazard in Banking, and Prudential Regulation: Are Capital Requirements Enough?" *The American Economic Review*, Vol. 90, No. 1 (Mar., 2000), 147 – 165. 作者十分感谢 Masahiko Aoki, Sudipto Bhattacharya, Jerry Caprio, Serdar Dinç, John Roberts, Paul Romer，两位匿名审稿人以及联邦储备银行纽约、曼海姆大学、斯坦福大学经济系和商学院，密执安大学 William Davidson 研究中心和世界银行参加讨论会的诸位参与者。本文存在的任何错误都由我们负责。本文的观点纯属是三位作者自己的看法，而与他们所在或者相关的机构无关。

历的银行危机，日本银行业的不良资产大约占到国民生产总值的1/4。[1]

审慎的监管措施的目的本来是保护银行业免受这些危机的冲击。一般而言，监管措施包括监督个人交易行为（比如保证贷款人有足够的抵押金），监管自我交易行为，制定资本要求（capital requirement），以及实施市场准入制度。在某些国家，政府会对于特定领域的借贷行为采取管制：比如许多亚洲国家都对房地产行业的借贷行为进行严格控制。[2] 最后，许多国家都实行利率控制制度。[3] 鉴于对银行挤兑的担忧，许多国家都采取措施实施存款保险，并且建立中央银行作为银行背后的银行。

在过去的10年中，原有的审慎监管体系发生了一些变化。首先，在交易的数量以及复杂性都上升的情况下，人们更加注重对银行风险管理体系的监控，而较少强调对个人交易行为的监督；其次，在一股金融市场自由化浪潮冲击之下，利率管制逐渐放开，而对银行选择资产组合的限制也更加放宽。第三，监管更加强调资本要求，基本上采用了巴塞尔国际银行协议（BIS）的标准。

尽管人们不断地改进监管体系，但是金融危机的爆发却越来越频繁（Gerard Caprio，Jr. and Daniela Klingebiel，1997；Klaus P. Fischer and Martin Chénard，1997）。许多现象都证实了道德风险所起到的重要作用。爱德华·凯恩（Edward Kane，1989），李贝尔·A. 科尔等人（Rebel A. Cole et al.，1995）为这个问题的注解为“为了明日复兴而投机”：银行选择具有风险的资产组合，其结果要么是大获全胜，从中获得高额利润或者红利，要么是一败涂地，给存款的储户或者说它们的保险人带来巨大损失。[4]有一种观点

1 基本上，不良贷款的处置率都在50%以下，如果日本的情况也是如此的话，它的损失可能会超过国民生产总值的10%。

2 这些措施的目的并不仅仅是确保银行体系的安全性和良好运行，还是为了将资本导入在当时看来回报更加丰厚的投资领域。

3 控制利率变化与许多其他的规制措施一样，还为其他的目标服务，对于国民生产总值负债率较高的政府而言，比如第二次世界大战之后的许多国家，较低的利率是具有重要的正面效应的。我们在Hellmann and Murdock（1997）以及Hellmann et al.（1997，1998a）讨论了这个问题。

4 乔治·A. 阿克洛夫和保罗·M. 罗默（George A. Akerlof and Paul M. Romer，1993）对于道德风险的问题有着更加深入的讨论，他们提出银行完全可以采取欺诈手段操纵贷款交易（比如内部贷款）来“劫掠”银行。在这种情况下，银行经理会榨干银行的所有价值，哪怕这样会导致银行破产。

认为，对存款进行保险是问题的根源，因为它降低了储蓄者监督银行的激励。[5] 而另外的观点认为，一个国家是否采取正规的存款保险制度是无关紧要的，因为一旦发生了金融危机，政府都绝不会袖手旁观。[6] 无论是具有正规的存款保险制度的国家，还是不具有正规的存款保险制度的国家都曾经爆发过金融危机，因此，仅仅依靠正规的存款保险制度本身是无法解决这个问题的。

本文试图探寻金融自由化本身是否是导致此问题产生的部分原因。近来的一些经验研究表明，这是有可能的：伴随着金融市场自由化，利率的巨大增幅是与金融危机的爆发存在着系统相关性的（Asli Demirgüç – Kunt 和 Enrica Detragiache，1997，1998）。本文从一个纯理论的角度提出这个问题。金融市场的自由化导致了竞争程度的加剧，竞争程度的加剧蚕食了银行利润，而较低的银行利润则降低了银行特许权的价值（即未来利润值资本化之后的期望值），而较低的银行特许权的价值又降低了银行保证贷款质量的激励，从而加剧了道德风险的问题。当竞争激烈程度达到一定限度之后，银行就会有投机的意愿。因此，在利率自由化和银行的审慎经营行为之间存在着不一致性。

我们想知道增加资本限制条件是否能够抵消自由化的负面效应。如果银行的资本金足够充足，它们完全可以将投机的负面后果内生化，从而采取更加审慎的投资行为。尽管通过设置资本要求能够解决道德风险问题，但是我们发现，它要求银行必须持有不必要的高额资本金。因此，仅仅依靠通过设置资本要求作为监管银行审慎经营的手段，是无法实现帕累托有效的结果的。也就是说，*存款利率自由决定与帕累托有效之间是不一致的*。

我们接下来想知道是否存在着能够使得均衡结果位于帕累托边界上的方法。仅仅依靠通过设置资本要求是不够的原因在于在存款利率自由决定的条件下，银行受到过高的通过提高存款利率来吸引存款的激励。只有当资本要求将银行成本提高到足以影响银行支付高额的存款利率的意愿时，资本要求才是有效的。既然要把资本要求作为降低存款利率的间接手段，那么为什么

5　然而人们可能会质疑监督银行存款的有效性以及意愿程度，只要监督银行存款是一项公共产品。Stiglitz（1985，1992，1994）以及我们的工作论文的讨论（Hellmann et al.，1998b）。

6　正如某位审稿人妙语所言，世界上一共有两种国家：那些具有正规的存款保险制度的国家以及那些还不知道自己具有正规的存款保险制度的国家。

不直接控制存款利率呢？实际上，我们会证明通过将资本要求和存款利率控制结合起来可以实现任何帕累托有效的结果。

似乎从一开始监管者就了解到，实施存款利率控制的好处就在于可以提高银行系统的稳定性。铃木淑夫（（Yoshio Suzuki，1987，p. 41）在评论1947年《临时利率调整法案》（“Temporary Interest Rate Adjustment Law”）的时候，曾经说道：

“制定这项法案的目的理所当然是为了避免利率竞争，因为后者会毁掉所有金融机构的利润。也就是说，相比通过内部控制确保银行审慎的经营行为而言，此项法案更加注重通过外部的存款利率控制来为各类金融机构营造一个稳定的经营环境。”

同样，美国在经历了储蓄与信贷危机之后，管制措施的一个主要变化就是废除了控制存款利率的第Q项条例。迈克尔·C.基里（Michael C. Keeley，1990）发现美国在整个20世纪80年代，在竞争加剧、银行专有权价值降低以及银行倒闭数目之间是存在着直接关系的。

有一种观点认为，金融市场的自由化至少应该对最近爆发的东亚金融危机以及日本脆弱的金融体系负一部分责任。[7] 首先，金融市场降低了市场的进入门槛，从而加剧了竞争：更多的外国银行进入国内市场；对于外国开设分支机构的限制也降低了；那些原本设立了存款利率上限制度的地区要么将存款利率上限的限制逐渐弱化，要么干脆完全取消。所有这些措施的后果就是利润率的下滑，以及随之而来国内现有银行专有权价值的降低。其次，自由化进程的其他方面降低了对银行施加的管制。原本禁止的一些经营行为例如，许多衍生产品交易和外汇交易也逐渐被认可了，于是这为银行参与投机经营开辟了许多崭新的渠道。更为重要的是，对房地产行业借贷行为的限制被取消了。[8] 再次，金融市场自由化在损害银行特许权价值并且为其开拓新的投机机会的同时，却没有采取相应的措施来加强监管。一方面资本要求反而被放宽，而其他的审慎控制手段又被取消了。因此，事实上监管机构监管

---

7 关于1997年东亚金融危机背后更加深入的动态机理请参见 Steven Radelet and Jeffrey D. Sachs（1998）以及 James Tobin（1998）。

8 房地产行业的借贷行为使得银行承担了大量的风险，这是因为市场上的房地产价格的波动性是非常大的。其程度如此之大，以至于人们很难确定房地产的真正价值，金融机构还会将资本的价值变得非常复杂（尤其是当银行的资产不必通过市场会计审核的时候），因为银行可以通过出售增值的资产，而持有贬值的资产来隐瞒真实资产状况。

市场的能力经常是不升反降的。[9]

在我们的模型之中，我们检验了在动态条件下的银行道德风险问题。银行可以通过投资风险小的资产来谋取高额预期收益，也可以投资结果不可预测的投机资产，如果投机一旦成功，银行就能够获取巨额私人收益，不过，如果一旦失败，就会给存款人带来巨额损失。如果金融市场的竞争足够激烈，那么银行通过审慎经营所获取的收益是非常有限的，但是银行总是能够通过投机抓住某一期的租。因此加剧竞争只会增加银行产业的投机程度。

这样，就需要某种审慎的管制措施来引导银行审慎投资。资本的要求使得银行将更多的资产置于风险之下，使得银行将投机风险内生化了。显然，一旦银行将相当部分的自有资产进行投资，它就会尽量从事审慎的投资。本文提出了一个审慎管制的替代形式：通过采取存款利率上限的限制措施来创造银行特许权的价值。银行特许权的价值是指银行未来利润流的现值，银行只有在不倒闭的情况下才能够获得银行特许权的价值。如果银行投机失败从而倒闭，那么它就会失去银行特许权的价值。因此，银行特许权的价值好像就是无形资产，但是可以作为现实资产的替代品。如果银行具有足够大的银行特许权的价值，那么它就一定会审慎地投资。

假定两种审慎管制形式可能是有效的，本文希望能够确定最优的管制形式。在每个存款利率水平条件下所需要的最低资本水平就形成了帕累托边界，这与银行选择进行审慎的投资是一致的。

为了说明为何银行自由决定存款利率能够导致帕累托有效的结果，考虑当某银行的所有其他竞争对手都选择了帕累托有效的存款利率的情况。在帕累托边界上，该银行对于投机冒险或者审慎投资两者是无差异的（就是说同等数量的存款所获得的银行利润是完全相同的）。那么，该银行投机冒险所获得的期望利润水平一定是高于审慎投资所获得的期望利润水平，因为如果该银行选择投机，一旦失败，银行就会损失掉特许权价值。如果该银行也选择了与所有其他竞争对手都相同的存款利率，那么就一定会审慎投资获取相应的市场份额。但是，如果该银行稍微提高一点存款利率的话，那么就能

9 例如，在泰国（别的国家也是一样的），私人部门的工资水平增长很快，连同公共政策当局的财政约束一起导致了公共部门和私人部门的工资水平不匹配，从而导致中央银行和政府部门流失了许多优秀的人才，他们都去了私人部门。

够吸引全部的额外存款，而如果它选择拿所有这些存款去投机，就会获得更高的利润水平。由于存在着窃取市场的效应，在帕累托边界上的任意位置上，每一个银行都受到偏离原有均衡的激励。于是，我们发现市场自由决定的存款利率不是帕累托有效的结果。

本文对此提出的对策，主要就是通过控制存款利率，创造出银行的特许权价值。通过排除银行利用无效率的高存款利率竞争，存款利率上限政策可以有效地削弱窃取市场的效应（market—stealing effect）。我们证明了在存款利率上限的条件下，可以生成任意的帕累托有效的结果。

我们的结论表明，在市场自由化与道德风险难题之间存在着清晰的理论联系。我们发现市场自由决定的存款利率不是帕累托有效的结果。银行采取无效率的高存款利率，以便从竞争对手那里挖来更多的客户。在通常情况下，市场自由化并不是旨在加剧金融市场竞争的程度。但是它的确存在着能够提高银行之间存款弹性的效应，而它对于总体存款弹性的效应，就算存在的话，也是极为有限的。这种银行之间贷款弹性的增加，会扩大窃取市场的效应，正是如此，才使得自由化与金融犯罪之间存在着必然联系。

当然能够改进单一资本要求的措施并不仅仅限于本文所详细研究的情况（存款利率上限）。政府能够采取的其他手段包括资产级别限制、市场进入许可限制，以及加强直接监控。这些措施的目的要么是为了限制银行介入道德风险问题的能力，要么是为了提高银行从事合意的审慎投资的激励水平。计算出所有这些措施的最优的组合方式已经超出了本文的范围。本文的要旨首先在于证明只靠资本要求是不够的，其次在于确认那些包括存款利率上限政策在内可能是有效的审慎管制措施。

在以前的文章之中，我们曾经探讨过其中一些更加广义的政策问题（Hellmann et al.，1996，1997，1998a；Hellmann and Murdock，1997）。在那些文章之中，我们考虑了一个金融市场管制的框架。我们提出了一系列称之为“金融约束”的政策建议，我们的这些政策建议与金融抑制政策是完全不同的。其中一个重要的区别在于，在金融约束条件下，利率控制完全是为了提高私人金融市场的斜率而制定的，而在金融抑制条件下，利率控制完全是政府为了从私人部门抽取租金的一种机制。

本文的理论模型是以苏迪普托·巴特查亚（Sudipto Bhattacharya，1982）的论文为基础，巴塔查亚注意到控制存款利率在简单的静态条件下

的用处。[10]卡普里奥和劳伦斯．H. 萨默斯（ Caprio and Lawrence H. Summers, 1996）强调了银行特许权价值对于审慎管制的重要性。[11]本文关于资本要求的讨论参考了让－克劳德·罗切特（Jean－Claude Rochet, 1992）的文章，他解释了资本要求是怎样影响银行投机的激励的。[12]在分析的过程中，我们吸取了其他人关于资本要求的看法，比如他无法确认所有相关的风险（Mathias Dewatripont and and Jean Tirole, 1994）。[13]

本文剩余部分的安排如下：第一部分给出了模型。我们接着在第二部分分析了完全竞争均衡并且导出了在均衡条件下投机行为产生的条件。第三部分考察了审慎管制的手段并且着重分析了运用资本要求和控制存款利率之间的权衡关系。第四部分介绍了紧的和松弛的存款利率控制，并且证明了松弛的存款利率上限约束是如何通过限制均衡路径下银行的投机行为来促进银行审慎经营的。在第五部分，我们考察了引入收益率内生化所产生的回馈效应，后者加强了我们从简单模型之中得到的结论。第六部分是结论。

一、模型

考虑一个生存期为 T 的银行。在其中的每一期，银行提供的存款利率为 r_i，而所有其他银行提供的存款利率为 r_{-i}。该银行能够吸纳到的所有存款量为 $D(r_i, r_{-i})$，存款量的大小随着本银行存款利率的上升而增加，随着其他银行存款利率的上升而下降（$D_1>0$，$D_2<0$）。[14]

由于存款人对存款实际上是进行了保险，因此存款量的大小完全取决于

10 布鲁斯·D. 史密斯（Bruce D. Smith, 1984）也建立了一个模型，其中控制存款利率可以降低金融波动的可能性。在他的模型之中，控制存款利率根除金融不稳定性的方式是与 Michael Rothschild and Stiglitz（1976）相同的。

11 Keeley（1990），Steven R. Weisbrod et al.（1992）以及 Rebecca S. Demsetz（1996）为此提供了关于美国和日本的经验证据。

12 Chun H. Lam and Andrew H. Chen (1985), Gerard Genotte and David Pyle (1991) 以及 David Besanko and George Katanas (1996) 证明了在某些条件下，资本要求可能真的影响到资产组合的风险。

13 对此的批评如下：（i）在调整过程之中存在着刚性问题；（ii）基本上这些模型的分析方法都是采取逐个考察资产的风险进行分析，这样就忽视了资产之间的联系。而且（iii）模型重点考察了信贷风险，但是常常忽略了市场风险。

14 需求函数需要满足的唯一额外条件就是它必须满足银行价值函数的凹性假设。

银行所提供的存款利率。[15]在我们看来，存款保险是最符合实际的假设条件。不过值得注意的是，我们的结论并不依赖于这个假设条件。在我们的工作论文那个版本之中（Hellmann et al.，1998b），证明了在没有存款保险的条件下，所有这些定性结果都是成立的。

当募集到足够的资金之后，银行会把它的资产进行优化配置，但是银行在决定贷款比例的时候面临着严重的道德风险难题。为简单起见，假设银行只能在两种资产之间选择：收益为 α 的安全的资产，以及以 θ 的概率获得收益 γ 和以 $1-\theta$ 的概率获得收益 β 的投机资产。安全的资产其期望收益水平较高（$\alpha>\theta\gamma+(1-\theta)\beta$），但是如果投机获得成功，银行就能够获得较高的私人收益（$\gamma>\alpha$）。银行的投资既包括募集到的资本，也包括银行自己拥有的资本 k，它被表示为募集资本的一个比例，这样所有的投资水平就等于 $(1+k)D(r_i, r_{-i})$。

资本的机会成本等于 ρ。[16]本部分我们假设 ρ 是外生的并且 $\rho>\alpha$（银行资本是有成本的）。简单的显示偏好原理表明 $\rho>\alpha$ 的情况是符合实际的。如果资本真的不存在机会成本，那么道德风险的难题也就不会像这样困扰整个现代银行业了，而银行也会很心甘情愿地合作。在第五部分我们将会更加深入地探讨这个问题。我们通过将 ρ 内生化扩展了这个模型，前者由使银行资产市场出清时的均衡资本收益率水平决定。在不存在资本要求的情况下资本是没有成本的（$\rho=\alpha$），但是银行实际上会持有很少的资本，甚至会不持有资本。不过紧的资本要求会增加银行对资本的需求，从而内生地产生了资本成本（$\rho>\alpha$）。[17]

通常银行要受到政府审慎管制措施的约束。在每一个时期的末尾，监管当局会派人督察所有银行的会计报表。如果某一家银行的净资产为负（即银行无法完全偿付所有存款人的存款），那么银行的特许经营权就被取消了。因此，如果银行参与了投机并且最终失败，那么银行就会丧失特许经营

15 尽管此结果对于所有固定费率保险都成立，但是我们的模型并不显含银行为存款保险所缴纳的费用。Yuk-Shee Chan et al.（1992）和 Ronald Gianmarino et al.（1993）都证明了怎样采取更加复杂的费用方案来降低道德风险问题。

16 我们可以把资本成本看作资本持有人资本被稀释的成本。股票投资者银行在每一个时期的期望收益（这也许取决于银行是否会参与投机，但是外部的股票投资者可以理性地预期银行的行为）。他们只有在投资的预期收益等于期望成本 ρ 时才会选择投资。

17 Gary Gorton and Andrew Winton（1997）在一般均衡条件下，同样地得到了银行资本有成本的结果。

权从而停止运营。[18]在这样的一个监管者（事后）具有完全信息的经济之中，是不会有那个净资产为负的“白痴”银行敢于冒风险“为了明日的复兴而投机”，指望搏一把成功之后还清所有欠款的（Kane，1989）。但是监管当局是无法在事前完全监督银行的投资资产组合的。正是如此，使得监管当局的监管策略从监督每一笔私人金融交易逐渐转换到监督银行风险资产结构上来。

银行选择安全资产的每一期利润水平为 $\pi_P(r_i, r_{-i}, k)=m_P(r_i, k)D(r_i, r_{-i})$，其中 $m_P(r_i, k)=\alpha(1+k)-\rho k-r_i$ 是银行在除去资本成本之后每单位存款所获得的利润水平。如果银行选择了投机，每一期利润水平为

$\pi_G(r_i, r_{-i}, k)=m_G(r_i, k)D(r_i, r_{-i})$，其中

$m_G(r_i, k)=\theta(\gamma(1+k)-r_i)-\rho k$，投机条件下每单位存款所获得的利润水平取决于投机是否成功。如果成功了，那么银行就获得了很高的资产收益，从而可以偿付所有的存款人。不过如果银行失败的话，银行就会被监管当局关闭。

银行的目标是最大化其期望利润的现值 $V=\Sigma_{t=0}^{T}\delta^t\pi_t$ 遵循道格拉斯·W. 戴梦德（Douglas W. Diamond，1989）的方法，我们考察 $T\rightarrow\infty$ 的情况。银行所选择的策略必须符合无限期重复静态博弈的纳什均衡的结果。

这个多阶段博弈的时间结构如下：所有银行同时决定各自的资本水平以及存款利率。存款人然后选择将资产存放到某家银行之中。最后利润得以实现，监管当局督察了银行的资产会计报表。[19]

二、投机条件下的竞争均衡

在比较各种审慎管制措施的形式之前，有必要考察一下所有这些管制措施是否是必要的。尤其是本部分的任务是分析银行在均衡条件下进行投机的

18 这里假设如果银行投机失败，其收益是无法偿付存款人的。很容易证明当银行在均衡条件下选择了投机时，这个假设总是满足的，因为投机资产的预期收益要低于安全资产。只有当银行能够对存款人（更准确地说是存款保险人）征收成本的时候，投机的银行私人收益要更高一些。

19 我们必须注意到此模型在每一期都是静态的结构。已经有越来越多的文章（Charles W. Calomiris and Charles M. Kahn，1991；Mark J. Flannery，1994）开始采取更为复杂的每期投资都是动态的结构。在不存在存款保险的条件下，这种方法是非常有效的。因为那样的话，人们可以分析存款需求的作用以及存款人取款时间间隔能够起到的作用。此类文献一般都认为，存款人间隔取款可以在一定程度上缓解，但是无法根除道德风险难题。

条件。银行投资安全资产的期望收益为 V_P （r_i，r_{-1}，k） $=\pi_P$ （r_i，r_{-i}，k） ／（$1-\delta$）。而银行投资投机资产的期望收益为 V_G （r_i，r_{-i}，k） $=\pi_G$ （r_i，r_{-i}，k） ／（$1-\delta\theta$）。银行的投资过程一共分为两个阶段：吸纳存款阶段以及资产配置阶段。在资产配置阶段银行拥有 D（r，r_{-i}） 个单位的资本用于投资，其资本成本为 r。如果 V_P （r，r_{-i}，k） $\geqslant V_G$ （r，r_{-i}，k），银行就会选择投资安全资产。否则银行就会选择投资投机资产。通过这种关系我们就能够推导出“无投机条件”，他决定了市场允许投机行为存在的利率水平下限值，即

$$\pi_G(r,r_{-i},k)-\pi_P(r,r_{-i},k)\leqslant(1-\theta)\delta V_P(r,r_{-i,k})$$

这个条件是符合直觉的。银行选择投机，每一期所获得的期望收益 $(\pi_G-\pi_P)$ 一定是小于投机失败（其概率为 $1-\theta$）所丧失掉的特许经营权的价值的（δV_P）。由此我们得到关键的存款利率水平 $\hat{r}$（k），只要 $r\leqslant\hat{r}$（k），银行就会选择投资安全资产（假设存款利率是满足对称均衡条件的）：

$$\hat{r}(k)=(1-\delta)\left(\frac{\alpha-\theta\gamma}{1-\theta}\right)(1+k)+\delta[\alpha(1+k)-\rho k]$$

当银行的眼光非常长远（即 $\delta\to1$），银行可以选择某个存款利率使得其接近银行的资本净收益水平［α（$1+k$） $-\rho k$］，并且仍然选择投资安全资产。这样假设是有意义的，因为当 $\delta\to1$ 银行所在乎的仅仅是每一期的平均收益，因而银行不会为了得到某个有限正值的当期利润而卷入投机，从而丧失掉所有未来收益的。然而一旦银行不是完全的眼光长远，那么银行一定可以在每一期挣得大量的利润，因此银行的特许经营权价值就必然大于投机的期望收益值。[20]

在给定了银行的资本配置决定之后，我们再来考察银行是如何在竞争市场上吸纳到贷款的。我们假设如果存在着无投机的竞争均衡的话，那么银行一定是会选择该均衡水平的。如果银行倾向于投资安全资产，那么它就会选择

$$(r_P,k_P)=\arg\max_{r,k}\{V_P(r,r_{-i},k)\}$$

在对称均衡的情况下（即 $r_{-i}=r_P$） 利用一阶条件（$\partial V_P/\partial r_i=0$） 可以求出

20 剩下的一个无投机条件下比较静态分析是十分简单的 $\partial\hat{r}/\partial\alpha>0$，$\partial\hat{r}/\partial\gamma<0$，$\partial\hat{r}/\partial\theta<0$。当安全资产变得更加具有吸引力的时候，无投机约束条件就变得越发松弛；而当投机资产变得更加具有吸引力的时候，无投机约束条件就变得更紧。

$$m_P\ (r_P,\ k)\ =D\ (r_P,\ r_P)\ /\ (\partial D\ (r_P,\ r_P)\ /\partial r_i)$$

通过此式可以求出 $r_P\ (k)$ 的显式解。用式进行替代 $\varepsilon\equiv\ (\partial D/\partial r_i)\ (r/D)$，得到

$$r_P(k)\ =\ [\alpha(1+k)\ -\rho k]\varepsilon/(\varepsilon+1)$$

由于 $\partial V_P/\partial k=-\ (\rho-\alpha)\ D\ (r,\ r_{-i})\ /\ (1-\delta)\ <0$。银行增加资本持有量反而会降低预期利润水平，于是银行会将自身所持有的投资资本最小化，因此

$r_P\ (0)\ =\alpha\varepsilon/\ (\varepsilon+1)$。当存款的竞争加剧到一定程度的时候（即当 $\varepsilon\rightarrow\infty$），竞争性存款利率就达到 α 的水平，同时银行的特许经营权的价值也变得任意小。但是如前所述，一旦存款利率超过一定的临界值，银行从投机资产中获得的收益要大于从安全资产之中获得的收益 $[r_P\ (k)\ >\hat{r}\ (k)]$，因此，所有银行都只投资安全资本的均衡情况是不存在的。

如果我们将注意力集中到银行选择投机资产的情况，我们可以重复推导前面的逻辑来证明

$$m_G(r_G,k)\ =\ \theta D(r_G,r_G)/(\partial D(r_G,r_G)/\partial r_i)$$

这就意味着

$$r_G(k)\ =\ [\gamma(1+k)\ -\ (\rho k/\theta)\ ]\varepsilon/(\varepsilon+1)$$

同上，可得

$$\partial V_G/\partial k\ =-\ (\rho-\theta\gamma)D(r,r_{-i})/(1-\delta\theta)\ <0$$

因此没有银行会愿意持有资产。

上面的结论可以归纳为下面的命题。令 $\bar{\varepsilon}\equiv\hat{r}\ (0)\ /\ [\alpha-\hat{r}\ (0)]$

命题 1：在市场存在着充分竞争的情况下（即 $\varepsilon>\bar{\varepsilon}$），唯一的对称均衡就是所有银行都选择不持有资本，向存款者支付 $\gamma_G\ (0)$，并且选择投资于投机资产。

三、审慎监管

一旦市场存在着充分的竞争，在自由竞争的市场均衡之中就必然存在着投机，因此就需要引入某种监管手段。在经济学家和政策制定者之间就以下观点已经达成了共识：银行最低资本要求与中央的有效监管共同组成了对银行审慎监管的有效手段。实行银行最低资本要求的原因在于，一旦银行对自有资本进行投资，这样自有资本就扮演了债券的角色，银行就承担了投机资产的一部分风险（我们将其称为风险资产效应）。只要银行拥有足够多的风险资产，银行一定会在均衡条件下选择安全资产。

还有一种替代性的审慎管制的手段就是控制存款利率，为银行支付给存款人的利率设置上限。通过限制存款市场的竞争，控制存款利率能够增加每个银行在每一期的利润水平，因此就增加了银行的特许经营权的价值。正如第二部分讨论的，一旦风险条件下的银行特许经营权的价值超过了投机资本的当期收益，银行就不会选择投机资产了。

为了有效地实施资本限制条件的政策（竞争均衡之中不存在投机行为），均衡条件下银行偏离均衡行为，而选择投机一定不是有利可图的。如果要求银行必须持有 k 的自有资本，假设所有的银行都选择投资安全资产，那么均衡的存款利率就是 r_P（k）。如果银行想要偏离均衡行为，那么它一定会选择能够使得收益达到最大化的存款利率水平，条件是只要其他的银行都提供 r_P（k）水平的存款利率，并且只有该银行投资了投机资产。为了使得这种偏离行为是无利可图的，从而确保数量为 k 的自有资本能够实现无投机行为的均衡，就必须要求下式成立：$\mathrm{Max}_r\ \{V_G(r, r_P, k)\} \leqslant V_P(r_P, r_P, k)$。[21] $\bar{k}$ 为满足该条件的最低资本。我们知道 $\bar{k}$ 总是存在的，因为对于足够大的资本量水平 k，银行所承担的风险成本会大到使得投机的收益严格小于投资安全资产的收益。由此我们还可以求得最低资本要求的比较静态结果：

$$\partial\bar{k}/\partial\alpha < 0$$
$$\partial\bar{k}/\partial\gamma > 0$$
$$\partial\bar{k}/\partial\theta > 0$$

由于金融市场自由化的浪潮，银行在决定贷款资产组合时有了更大的自由度，我们可以想像 γ 或者 θ（或者两者都会）的水平会上升。从而由于金融市场自由化而造成的最低资本要求提高的可能性是增加了。[22]

实施有效的存款利率控制这一政策是相对简单的。因为对于所有 $r \leqslant \hat{r}$（k）的情况，银行投资安全资产是最优的选择，因此任何 $\hat{r}$（k）水平的存款利率结合数量为 k 的最低资本要求政策，将会实现无投机的均

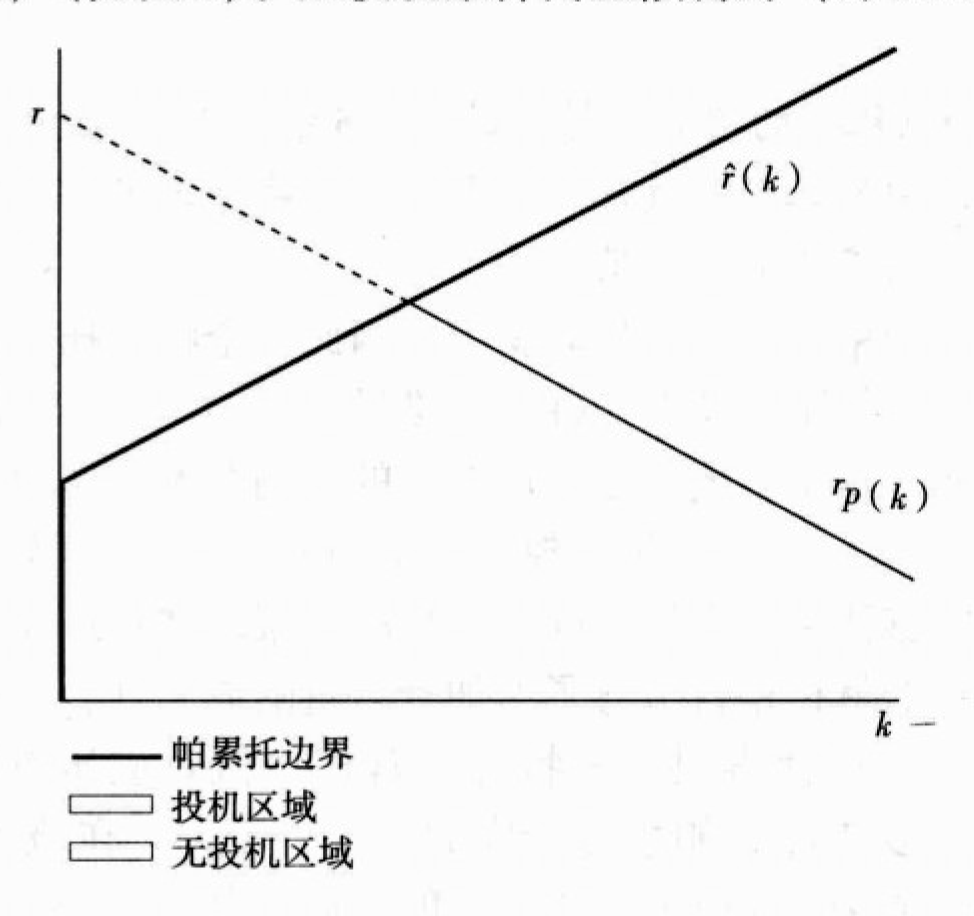
21　为了标记的方便，当书写 $r_P(k)$，$r_G(k)$ 和 $\hat{r}(k)$ 作为函数形式的时候，我们会略去他们都是银行资本水平的函数的事实，并且写作 r_P，r_G 和 $\hat{r}[V_P(r_P, r_P, \bar{k}) = V_P(r_P(\bar{k}), r_P(\bar{k}), \bar{k})]$。

22　这可能是一个很重要现象，《巴塞尔资本协议》的目标之一就是为国际竞争创造一个平等的“竞技场”。但是如果不同国家的银行所面临的投机机会和竞争激烈程度都不同，那么不同国家最优的资本要求也是不同的。

衡。[23]并且即使不存在资本要求，$\hat{r}$ (0) 的存款利率上限水平就足以确保结果的有效了。我们把这个结果总结为命题2。

命题2：$\hat{r}$ (k^*) 水平的存款利率结合数量为 k^* 最低资本要求的政策，将会实现任何帕累托有效的均衡结果。

银行、存款人以及政府（作为存款的保险方）的收入水平是存款利率，银行持有的资本量水平以及银行到底是审慎投资还是投机的函数。这样帕累托边界就取决于以下一系列结果：没有任何代理人（银行、存款人以及政府）能够在不损害其他人的利益的情况下提高自身的福利。在无投机的区域内帕累托边界的第一个部分是由 $\hat{r}$ (0) 以下的纵轴定义的（所有 $r \leq \hat{r}$ (0)的情况）（如图1）。无投机条件向上倾斜的（即 $\partial\hat{r}/\partial k > 0$）。



注：$\hat{r}$ (k) 线是无投机条件，r_P (k) 线是审慎投资的均衡存款利率条件。r_p (k) 线高于 $\hat{r}$ (k) 线是无法实现的，这是因为在这个区域，银行绝对是愿意投机而不是审慎投资。

图1 无投机区域与竞争性均衡

但是在检验无投机条件的时候，我们发现它并不一定是向上倾斜的：

$$\partial\hat{r}/\partial k = (1-\delta)(\alpha-\theta\gamma)/(1-\theta) - \delta(\rho-\alpha)$$

23 名义上我们在实施控制存款利率水平时只考虑存款上限。如果 $k^* > \underline{k}$，其中 $\underline{k} = \{r_P(k) = \hat{r}(k)\}$，那么在存在存款利率上限但是不封底的竞争均衡条件下，银行会选择低于的存款利率水平。为了实现该条件下的社会最优化，监管当局也必须实行存款利率封底 $\hat{r}$ (k^*) 结合数量为 k^* 的资本要求的政策。

就是说，对于所有 $\delta \geqslant \bar{\delta} \equiv (\alpha-\theta\gamma) / [\alpha-\theta\gamma+(1-\theta)(\rho-\alpha)]$。有 $\partial\hat{r}/\partial k \leqslant 0$。这个发现就意味着对于目光远大的银行而言，增加银行自有资本存量实际上是增加了银行进行投机的激励。这个结果是与传统文献所发现的结论完全相左的（Bhattacharya，1982）。传统的分析方法往往注重银行资本的静态作用。即银行资本金之中自有资产比例越高，银行内生化其投机成本的程度越大。但是这种静态分析方法没有考虑资本要求对于银行特许经营权价值的动态影响。因为持有资本是有成本的，如果银行增加资本金之中自有资产的比例，在其他情况不变的条件下，就降低了银行以后每一期所赚取的利润水平。因此银行增加资本金之中自有资产的比例具有两种效应：正的资本风险效应以及负的特许经营权价值效应。这个结论总结为下面的命题 3：

命题 3：如果银行的视界足够长的话（$\bar{\delta} \leqslant \delta < 1$），负的特许经营权价值效应就会压倒正的资本风险效应。于是，总是存在着某个存款利率控制政策帕累托占优于其他资本要求政策。

因为无投机约束条件是向下倾斜的，因此这个帕累托占优的结果可以由银行具有足够长远的眼光这一条件直接推导出来的。对于任何的正的资本要求，均衡存款利率水平都要低于采用单一的控制存款利率的手段得到的利率水平（即 $\hat{r}(k) \leqslant \hat{r}(0)$）。如果控制存款利率水平为 $\hat{r}(k)$ 即使没有资本要求，也同样能够让存款人得到同等水平的回报，并且还能使得银行获得更高的利润水平，这就让银行节省了增加资本量的成本（$[\rho-\alpha]k$）。只要银行的眼光长远到足以使得银行的特许经营权价值效应压倒风险资本效应，那么任意的资本要求都是帕累托最优的政策。于是，在该条件下最优的资本要求都是零。在本文的剩余部分，我们假设银行的眼光都是足够长远的，从而非投机条件是向上倾斜的（即 $\partial\hat{r}/\partial k > 0$），而最优的资本要求是有可能为正的。即使在非投机条件是向下倾斜的情况，我们的结果也是成立的。

我们已经证明了，总是可以通过采用资本要求结合控制存款利率水平，来实现约束条件下的最优均衡。现在，我们考虑是否能够在不采取控制存款利率只依靠资本要求的情况下——这正是金融自由化所经常采取的政策——实现最优均衡。

当不存在存款利率限制的时候，银行可以自由选择存款利率水平。为了理解仅仅采用资本要求的有效性，理解银行为何存在提供更高水平的存款利率并且将其信贷资产转向投机资产的激励是十分重要的。$\underline{k}$ 由

$r_P(\underline{k}) = \hat{r}(\underline{k})$式确定。由在$\underline{k}$的水平上银行具有偏离均衡提供更高水平的存款利率，并且将其信贷资产转向投机资产的激励的条件，可以推导出引理1，引理的证明详见本文附录。

引理1：*最低水平的资本要求是严格地大于$\underline{k}$（即$\bar{k} > \underline{k}$）。*

引理1的经济学直觉来源于银行非投机条件的实现方式与均衡条件下存款利率的实现方式的差异。非投机条件的实现方式是通过去除银行在资产决策时期的投机激励来实现的。因此在银行的存款水平是一定的条件下，银行依据其收益水平来选择是投资安全资产还是投机。如果银行对于投资安全资产还是投机（这正是$\hat{r}$的定义）是无差异的，那么银行投资投机资产所得到的当期期望收益一定是较高的：

$$m_G = m_P(1 - \delta\theta)/(1 - \delta) > m_P$$

我们接着考虑银行吸纳存款的时期，此时如果银行是想要投机而不是投资的，那么它就会有更高的吸纳存款的激励，这是因为吸纳的存款量越多，银行的利润就越高。

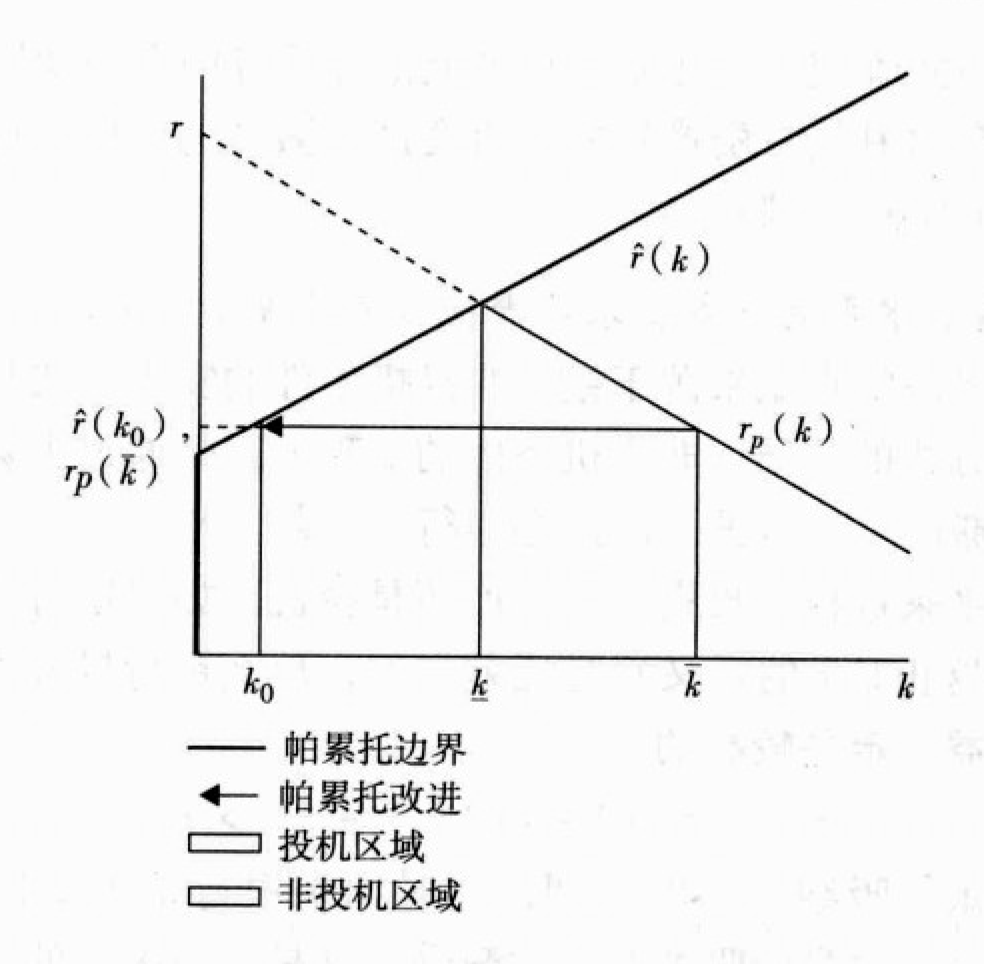
对于任何大于$\underline{k}$的资本要求，均衡条件下的存款利率水平一定是严格位于帕累托边界之内的，因为$r_P(k)$是向下倾斜的。这就意味着还存在着某个更低的资本要求k_0使得$\hat{r}(k_0) = r_P(\bar{k})$（见图2）。因此作为水平为$\bar{k}$的资本要求的一个替代，我们可以采取将$\hat{r}(k_0)$的存款利率限制结合$k_0$的资本要求的政策。这种替代政策对于存款人的报酬是相同的，但是能够给银行带来更高的利润水平。命题4总结了这个结果。

命题4：*总是存在着某个由资本要求以及存款利率控制措施组成的政策组合，能够帕累托占优于任何的单一的资本要求的。*

这个结论是很可靠的。它表明世界上的大多数国家所采取的政策体制（采用单一的资本要求而不去控制存款利率的政策）都不是帕累托有效的政策措施。

只要正的资本要求是引致银行采取审慎的经营行为的必要条件，这种政策体制就一定不是帕累托有效的。这个结论对于分析另一种创造银行特许经营权价值的替代方案是尤为重要的：实施市场进入许可制度。降低市场之中相互竞争的银行数目与降低银行相互之间的存款替代弹性是等效的。[24]这会反过来增加银行的特许经营权的价值。但是它还是不能实现帕累托有效的结

24 在霍太林模型（Hotelling model）之中可以推导出这个结果的显式解。



注：$\hat{r}$（k）线是非投机条件，r_P（k）线是对审慎投资的均衡存款利率条件，且无存款利率上限。存款利率上限或者 $\bar{r}=r_P$（$\bar{k}$）产生的结果是相同的，它使得政府能够将资本要求从 $\bar{k}$ 降为 k_0，而同时促使银行实行审慎投资。

图 2　有存款利率控制的帕累托改进

果。只要银行能够自由地决定存款利率水平，每一个都有沿着帕累托边界提高存款利率的正向激励。由于存在的市场进入限制不能影响银行制定存款利率的行为，激励的问题使得无法实现帕累托有效的结果。

有一个相关的问题就是，我们到目前为止都假设经济之中的银行数目是外生给定的。显然，在进入市场的决策是内生的条件下，银行的特许经营权的价值很可能受到潜在的进入者的蚕食。但是新的银行只有在沉没成本小于或等于事后的银行的特许经营权价值条件下，才会选择进入市场。如果沉没成本超过了审慎经营得到的特许经营权价值，那么进入市场的内生化就不会影响我们的结果。如果沉没成本太小，只要政府征收一个许可证费用，使得总的沉没成本（进入成本加上许可证费用）等于其特许经营权价值，那么结果依然成立。

最后，考虑一下没有资本要求仅仅实施存款利率控制措施的情况也是很有意思的。显然控制存款利率只能实现 $k=0$ 以及 $r\leqslant\hat{r}$（0）的结果。所以这些结果都是帕累托有效的，但是它们仅仅是帕累托边界上的一部分。

四、紧的以及松弛的存款利率上限

前面的结果描述了，采用存款利率上限的作为有效的审慎监管机制的优点。美国在20世纪60年代以及70年代所经历的事实说明了现实之中实施存款利率上限的政策存在着许多潜在的困难，而这些在我们的模型之中尚未正式地加以分析。尤其是当银行被剥夺了自由选择存款利率的权利的条件下，它们可能会采取其他的非价格竞争手段，比如分发消费品（烤面包机），开设新的分支机构，以及发布广告等来吸纳存款。这种非价格竞争手段是无效的，因为它使得消费者实施了无效的商品替代，并且它还可能会产生负面影响，使得由存款利率上限所创造的银行特许经营权价值最终被耗尽。[25]

然而，认为非价格竞争手段是社会浪费的观点不能不说是非常主观的。尤其是当金融部门的发展还处于不发达的阶段，非价格竞争手段是具有深化金融部门改革的正面效应的。在赫尔曼等人（1996）的文章之中，我们考虑了经济之中存在着服务不足的存款市场。银行需要首先进行某种投资（比如兴建服务点分支网络）来吸引客户。在这种情况下，非价格竞争手段是符合社会需要的，但是银行自身只有在存在着存款利率控制的条件下，才会有进行此类投资的激励。在这种情况下，金融深化的收益可能会超过由非价格竞争手段造成的无效性。[26]

当然，这些结论只是在发展中国家的条件下才适用的。在发达的工业化经济体中，金融服务的范围几乎遍及所有的家庭户，因而增设服务网点的数目所能够获得的收益是十分有限的。在这种情况下，存款利率控制可能会产生某种社会浪费性的非价格竞争，正如美国在20世纪60年代以及70年代所经历的那样。家庭户所能够拥有的金融产品越多（股票、债券，以及货币市场共同基金等），就意味着，家庭户越有可能在存款利率上限的政策条

---

25 然而非价格竞争不会最终耗散由存款利率上限所创造的银行特许经营权的价值。银行会采取非价格竞争手段直到其吸引额外一个单位的存款量的边际成本低于其边际收益。当边际成本是递增的条件下，银行就能够赚取所有的边际成本以下的存款的租。

26 同样地，在进入决策是内生化的条件下，非价格竞争也可能是有价值的。正如在前面论述的那样，新的银行只有在其进入的沉没成本小于或等于事后的银行的特许经营权价值时才会选择进入市场的。倘若银行进入的沉没成本会随着银行介入更多的非价格竞争而增加（例如，兴建新的分支机构采用昂贵的大理石材料），那么均衡条件下的银行的特许经营权价值会同样地增加。因此非价格竞争也可能对于促使银行决心投资安全资产具有正面作用。

件下，选择其他的金融产品来替代银行存款。

由于实施存款利率上限的审慎监管政策，这样就不可避免地会产生有关信息获取的相关顾虑。在这种政策体制条件下，存款利率水平是由政府制定的。然而政府在制定最优的存款利率水平的时候会面临许多困难，尤其是当宏观经济受到很大的冲击的时候。显然，这正是美国经济在20世纪70年代所经历的事实。条例为存款利率制定了上限。在随后的1973年石油危机冲击之下，这个存款利率水平低于通货膨胀率，这迫使银行向存款者支付负的真实收益率，从而存款人大都转向其他的金融资产。

即使在某些条件下，紧的（binding）存款利率上限不是有效的措施，我们还是相信存款利率上限的政策是进行审慎监管的有效手段。这是因为就算存款利率上限政策在均衡条件下是松弛的，也是十分有用的。松弛的（nonbinding）存款利率上限在不会影响投资安全资产的银行决策，但是会影响想要进行投机的银行决策。[27]因此松弛的存款利率上限十分吸引人，这是因为它不会引起无效的非价格竞争行为。此外，因为它允许监管者有更多的失误，它还能够降低监管者的信息问题。

为了在模型之中展示这一点，注意投机银行总是企图提供比审慎经营的银行更高的存款利率水平。当其他的银行都采取审慎经营策略的时候，偏离均衡准备投机的银行就会依照 $r_D(k) \equiv \arg\max_r \{V_G(r, r_P, k)\}$ 来制定存款利率水平。如果银行预期投机是有利可图的，即银行从投机资产上赚取的利润水平要高于从安全资产上赚取的利润水平（$m_G(k) > m_P(k)$），那么通过提供更高水平的存款利率 $r_D(k) > r_P(k)$ 就会使得银行偏离均衡。因此政府的确应该制定存款利率上限 $\bar{r}$ 以满足 $r_P(k) < \bar{r} < r_D(k)$，这样，该政策不会对均衡的存款利率水平造成影响，同时又能够降低投机的收益水平，因为（$V_G(\bar{r}, r_P, k) < V_G(r_D, r_P, k)$）。

这一点极为重要，因为（由引理1）正是投机银行具有的提高存款利率水平的超额激励，使得在不存在存款利率上限的情况下，政府不得不实施 $\bar{k} > \underline{k}$ 的资本量限制政策。倘若政府采取的存款利率上限为 $\bar{r}$ 并且对于任何 $\underline{k} < k \leq \bar{k}$ 都满足 $r_P(k) < \bar{r} \leq \hat{r}(k)$，那么所有银行都必须提供满足非投机条件的存款利率水平，并且银行偏离均衡进行投机是无利可图的。[28]这就有了

---

27 我们使用“松弛”这个术语来表示存款利率上限对于均衡条件下的存款利率水平所造成的效果。然而存款利率上限对于任何企图进行投机的银行而言是松弛的，这正是该条件有效性的根源。

28 注意对于所有 $\underline{k} < k \leq \bar{k}, r_P(k) < \hat{r}(k) < r_D(k)$，因此均衡条件下约束条件是松弛的。

命题5。

命题5：对于任何$\underline{k}<k\leqslant\bar{k}$，如果政府采取非投机存款利率上限为$\bar{r}\in(r_P(k),\hat{r}(k))$，并且均衡的存款利率水平由市场竞争决定$(r_P(k))$，并且所有的银行都采取审慎经营的策略。

由命题1到命题4，我们知道倘若不存在存款利率上限，政府就必须采取严格的资本要求，从而导致了帕累托无效的结果。之所以选择这个结果的一个最重要的理由就是，当政府仅仅采取资本要求的时候，存款利率是由市场力量决定的，并且（如前所述）由于我们的标准模型所没有阐述的原因，这是有价值的。命题5是对此论点的直接回应。如果均衡条件下的存款利率上限不是紧的，那么存款利率水平仍然由市场力量决定，但是我们就能够放松对于银行资本的约束。这样的结果就会比仅仅采取资本要求更加接近帕累托边界。

当然，政府仍然需要设定适当的存款利率上限。如果上限过低，约束就会变紧，这样就会导致以上所讨论的所有的结果；而如果上限过高，就无法阻止银行进行投机行为。[29]美国在实施Q条例的过程之中有一条经验是非常明显的，就是任何存款利率水平的上限必须是制定在真实水平上而不是名义水平上，这样，政策的效果就不会遭受期望通货膨胀率大幅波动的影响。我们可以把同等期限的国库券作为一种合理的衡量标尺。政府发行的国库券利率是由市场决定的，因此它很自然地包含了预期的通货膨胀率。如果存款利率上限被设定在高于同等期限的国库券利率的地方，那么该限额就将伴随市场情况的变化而不断地调整。[30]

还有一个原因使得控制存款利率上限超出同等期限的国库券利率的额度是有意义的：银行存款是向政府保险了的，因此银行实际上是在以政府的信用作为担保来借款。肖文等人（Shoven et al.，1992）认为20世纪80年代真实利率水平大幅突然地上升，是与储蓄和信贷危机的冲击相关的，从而导致在与同等期限的债权资产竞争的时候，银行存款不得不提高其利率水平。这里的关键之处在于当投了保的金融机构决定偏离原来的均衡而进行投机，他们在与政府竞争的时候，是能够利用政府信用来提供高额的利率水平的。

29　如果银行存在异质性，那么可能会使得找到合适的存款利率上限的范围更加复杂。显然资本要求也面临着同样的问题。

30　存款单据能够比同等期限的国库券提供更高的回报率。之所以会有如此大的差额，主要是因为两者在税收待遇方面的区别：存款的利息收入还需要缴纳地方税，而国库券的利息收入则不需要（Shoven et al.，1992）。

五、内生的收益率

截至目前为止，我们都将资本成本 ρ 和贷款资产的收益率 α 视为给定的。然而，可以设想在一般情况下，尽管每一个银行都将 ρ 和 α 视为给定的，但是在总量的水平上它们还是内生决定的。这就为政府决定干预这些内生变量提供了依据。本部分我们将引入该模型的一个简单扩展，它允许我们将银行资本成本和贷款资产的收益率内生化。其最主要的想法就是，对于银行资本成本的假设（ $\rho > \alpha$）完全是银行资本要求提高了银行对于稀缺的存款资本需求量的一个自然而然的结果。其次，内生决定这些收益率的反馈效应（feedback effect），倾向于放大存款要求，对银行进行审慎投资激励有负面影响。

我们都知道在一般均衡条件下可能会存在着多重反馈效应。这里我们所采取的方法是将分析局限于能够在许多合理的均衡模型之中都存在的那些一阶效应。特别地，我们在下面引入了总供给和总需求函数，它们具有一些天然的特性。对于这些总供给和总需求函数的微观假设条件，比如，对技术和偏好进行设定已经超出了本文的范围。相反，我们将采取“简化型”的方法，这样，我们就能够比较广泛地考察均衡回馈效应而不必依赖于某个特定的模型假设条件。

为了内生化 ρ 和 α，我们将对银行的资本金和银行借款分别采用总供给需求函数模型。首先考虑银行资本金市场，假设存款人能够自由选择储蓄的形式作为存款或者作为投资银行资本金市场的基金。我们稍稍滥用以下标记，将存款量的总供给标记为 D，而银行的资本金市场的基金总供给为 E。[31]许多一般均衡模型都采取的一个合理的假设条件就是基金的总供给函数是关于自身收益率（弱）递增的；并且关于其替代品收益率（弱）递减的。并且假设其自价格效应占优于其交叉价格效应似乎也是合理的。因此我们假设 $D(r,\rho)$ 和 $E(r,\rho)$ 满足

$D_r(\equiv \partial D/\partial r) \geqslant 0, D_\rho \leqslant 0, D_r + D_\rho \geqslant 0, E_\rho \geqslant 0, E_r \leqslant 0$，以及 $E_\rho + E_r \geqslant 0$。

两种总供给函数背后的思想在于银行资本金与其吸纳的存款之间是有着显著的差异的，而存款人可能会考虑这些差异。首先，银行为存款人提供了流动性需求以及诸如支票付款等其他相关的服务。其次，投资银行资本金可能会存在着交易成本，而且投资者参与市场投资的手续也可能十分

31 总供给函数是通过简单地将个人需求函数乘以银行数目 N。

复杂。最后这些证券的风险程度也可能会有所不同。尽管分析上的困难使得我们无法直接将这些效应模型化，我们的简化形式需求函数与以上的看法是一致的。

银行对于资本金的需求依赖于资本要求以及 ρ 和 α 之间的关系。将需求函数表达为存款量比例 $\tilde{k}$ 的函数是很方便的。令 k 为最低水平的资本要求。银行对于资本金的需求可以表达为 $\tilde{k}D$，其中当 $\rho<\alpha$ 时 $\tilde{k}=\infty$；当 $\rho=\alpha$ 时 $\tilde{k}\in[k,\infty)$；当 $\rho>\alpha$ 时 $\tilde{k}=k$。这就是说银行在 $\rho<\alpha$ 的条件下总是愿意提高其资本金至任何高水平的，而当 $\rho>\alpha$ 时，银行仅仅只想提高约束条件下的最低资本量。当 $\rho=\alpha$ 时银行对于提高多少资本量水平是无差异的。

银行资本金的供给和需求完全是符合标准的完全竞争市场模型，而银行以及存款人都将资本价格 ρ 视为给定的。供给和需求的交点决定了均衡水平。显然在 $\rho<\alpha$ 的条件下是无法构成均衡的，因为任意一家银行都可以提高资本回报率来吸纳更多的资金。为了了解 $\rho=\alpha$ 和 $\rho>\alpha$ 条件下是否可以构成均衡，定义 $k^V\geqslant 0$ 满足 $E(\alpha, r)=k^vD(r, \alpha)$。$k^v(r)$ 表示即使在资本要求不存在的情况下，银行自愿持有的资本量水平（上标 v 为正是自愿的意思）。[32]这里我们考虑银行自愿持有的资本量水平大到足以实现均衡条件下银行审慎经营的特殊情况。首先假设银行资本金的供给是非常充足的。这里特别考虑 $k^v(r_P(\bar{k}))\geqslant\bar{k}$ 的情况。在这种情况下银行自愿持有相当数量的资产并且不会参与投机。在这样一种经济的银行业中是不会存在道德风险难题的，而且实施审慎监管措施也是完全没有必要的。

不幸的是，现实世界与所描述的现象并不相符。我们发现一般情况下银行资本的回报率都必须补偿流动性的不足以及其他金融服务的欠缺。这就意味着银行资本金市场的供给水平很可能是不足的。特别是当 $k^v(r_P(\bar{k}))<\bar{k}$ 的情况下，银行自愿持有的资本量水平并不足以实现均衡条件下银行审慎经营。首先假设不存在资本要求。如果 $\rho>\alpha$ 那么银行不会愿意持有任何水平的资产量。还可以推导出在 $\rho=\alpha$ 条件下，为了去除均衡之中的投机行为银行所需要持有的资本量水平是下降的。当银行资本金市场出清，并且银行资产的回报率等于其资本成本，均衡条件下是存在着投机行为的。

正是在这种条件之下，监管当局想要实施某种资本要求。但是实施紧的

32 检验 k^v 是 r 的增函数是非常容易的，即使在 α 是内生的条件下也成立。

资本要求会提高其资本金需求，并且因此提高了资本金价格。对于最低的有效资本要求 $\bar{k}$ 其资本金市场在 $\rho > \alpha$ 的某一点达到均衡，使得 $E(\rho, r) = \bar{k}D(r, \rho)$。我们由此证明了高额的资本成本是可以由监管措施内生地引致出来。

对于 $E(\rho, r) = kD(r, \rho)$ 式求全导数，我们可以立即得到（采用直接的标记方法）$\rho_k = D/(kE_\rho - D_\rho) \geqslant 0$ 以及 $\rho_r = (kD_r - E_r)/(E_\rho - kD_\rho) \geqslant 0$。如果资本成本是内生的，那么它一定是 k 和 r 的增函数，更高的资本要求会提高银行对于资本的需求，这样就使得资本回报率上升。更高的存款利率水平同样会提高银行对于资本的需求水平，只有这样才能够与其高额的存款需求量相称。另外，更高的存款利率水平使得存款人投资银行资本金的决定更加困难，从而要求有更高的资本回报率，才能将存款人重新吸引到银行资本金市场上来。

我们还将模型扩展到了把贷款利率水平 α 的决定机制内生化的情况。同样地，定义一个完整的一般均衡模型超出了本文的范围，但是我们至少可以采用简化型的方法来分析主要价格对于数量的依赖关系（即贷款利率的水平与整个银行贷款总量之间的依赖关系）。[33]我们把对银行贷款的需求记为 $L(\alpha)$。大多数银行贷款模型都会预期贷款需求会是 α 的（弱）减函数（当银行增加其贷款利率水平或者收紧其他项目的时候（α 增加），企业可能就会削减其贷款需求）。而贷款的总供给是由银行全部资产提供的 $A = D + E$。信贷市场均衡由下列条件 $L(\alpha) = A(r, \rho(r,k))$ 给出，其中 $\rho(r,k)$ 的定义同上。对该式求全导数，结合 $L_\alpha < 0, A_r = D_r + E_r \geqslant 0$，以及 $A_\rho = D_\rho + E_\rho \geqslant 0$，可以得到 $\alpha_r = (A_r + A_\rho \rho_r)/L_\alpha \leqslant 0$ 以及 $\alpha_k = A_\rho \rho_k / L_\alpha \leqslant 0$，提高存款利率水平或者资本要求会降低银行贷款的回报率。

现在，我们就可以分析将 ρ 和 α 内生化是如何影响政策选择的。为此，我们将要再次考察非投机条件，后者定义了一系列决策者能够实施的 r 和 k 的可行政策组合。附录中给出了命题 6 的证明。

命题 6：*在扩展模型之中考虑提高资本要求，除了命题 3 推导出的常规的风险资产效应以及银行特许经营权价值效应之外，资本成本以及银行贷款回报率 α 的内生性意味着还存在着额外的"反馈"效应使得银行审慎经营的激励更加降低了。*

---

33 有一点是非常有趣的，就是贷款市场自身不断地运行会产生"银行特许经营权的价值"效应（"franchise-value" effect）。见 Serdar I. Dinç (1997)，他建立了这样的一个模型。

这种结果的经济学直觉是这样的。我们已经看到了提高资本要求会提高资本成本。而它将会进一步地降低银行特许经营权的价值，如此就损伤了银行进行审慎经营的激励。更重要的是，提高资本要求还会降低银行贷款的收益率，这会对银行特许经营权的价值效应以及静态的风险资产效应产生负面影响。[34]

图3 显示了该命题的含义。如果ρ和α是外生的，那么$\hat{r}$（k）就是一条直线。但是如果ρ和α是内生的，那么$\hat{r}$（k）的斜率就总是较低的，因此$\hat{r}$（k）的轨迹就会位于外生条件下的$\hat{r}$（k）直线之下。[35]这就意味着监管当局现在所面临的能够实现均衡条件下银行审慎经营的r和k的可行政策组合的范围变得更小了。我们还发现，如果银行的特许经营权的价值效应是占主导地位的，那么$\hat{r}$（k）的斜率就是负的。如果ρ和α是内生的，这种情况就更有可能发生，因为回馈效应也具有使得$\hat{r}$（k）的轨迹向右倾斜的效果。这就是说，如果回馈效应足够大的话，那么提高资本要求永远也不可能引致银行采取更加审慎的经营行为。

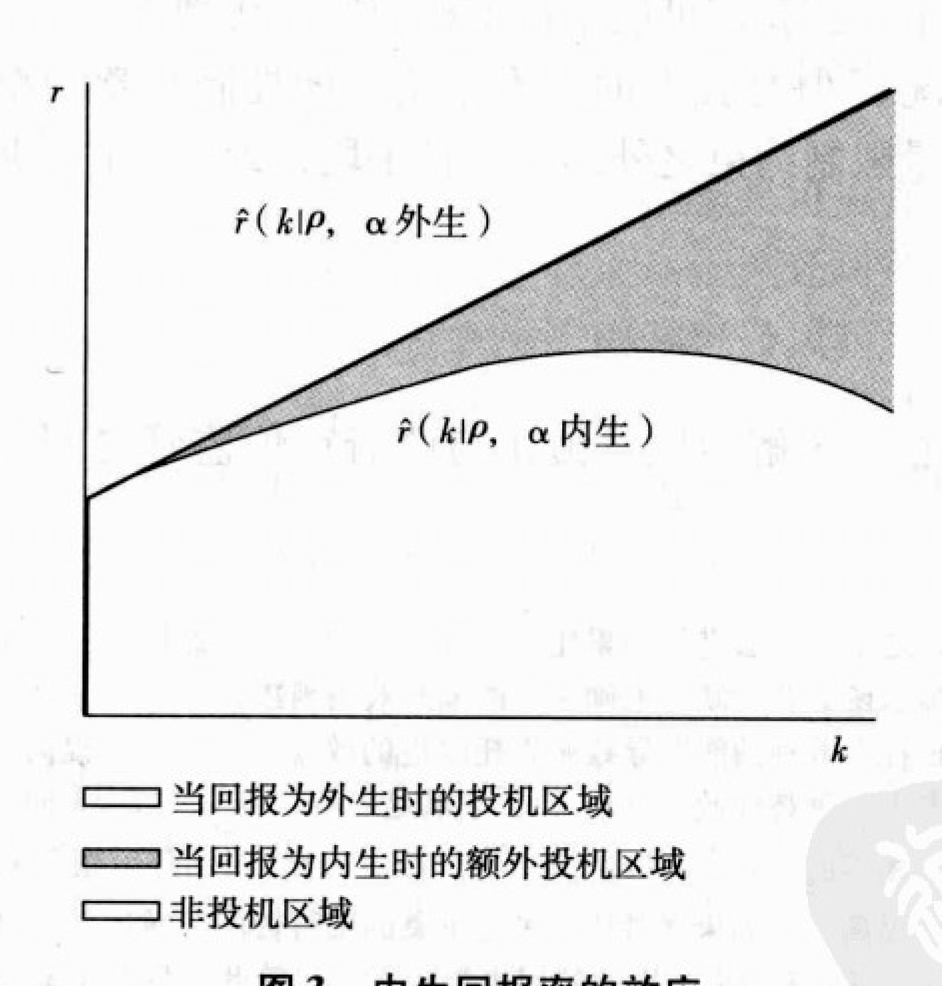


图3　内生回报率的效应

34　注意这里我们假设由θ、β和γ表示的投机的收益率是一个常数，只要在均衡条件下没有银行采取投机行为，那么这个假设就是最自然而然的。而且放松该假设也是很容易的。

35　外生的$\hat{r}$（k）线依照$\hat{r}$（0）水平上的ρ和α描绘出来。

有必要检查一下在本文第二部分回报率为内生的条件下，我们得到的结论是否正确。一般情况下我们对于消除道德风险难题所采取的资本要求可能造成的成本以及控制存款利率可能获得的利益的结论是很稳健的，但是我们关于帕累托有效性的分析假设必须放松。命题1和命题2可以由前面给出的条件直接推导出来。而命题6是命题3一般化。命题5的重点在于松弛的存款利率约束使得政府能够在降低最低资本要求的同时，保持存款利率由市场决定。在回报率是内生决定的条件下，放松资本要求具有额外的好处，因为这样银行的资本成本 ρ 也会下降。然而我们必须放松命题4的结论，因为我们再也不能够在扩展的模型之中进行正式的帕累托有效性分析了。这也并不令人吃惊：毕竟我们在模型之中引入了更多的代理人的类型，而代理人之间的差异越大，某一种政策就越有可能损害至少某一类型的代理人的利益。[36]举例来说，降低最小资本要求（如图2）会对银行贷款产生负面影响，于是贷款人的福利可能会下降。显然，降低最小资本量水平还会降低银行的资本成本（因为在均衡条件下有 $\rho > \alpha$）。[37]为了全面分析某一种政策的影响，我们可能不得不权衡所有代理人之间的福利分配这一种更为复杂的情况。[38]但是我们的基本结论，即控制存款利率水平为监管当局提供了除资本量限制约束之外的另一种手段，即使是在扩展模型之中也是成立的。

六、结论

本文的目的在于理解金融自由化与政府审慎监管之间的相互作用。金

---

36 从技术上讲，之所以无法进行帕累托有效性分析的另一个原因，是因为在扩展模型之中所使用的简化型供给和需求函数背后缺乏准确定义的偏好效用函数。

37 这里有可能还存在着别的能够导致帕累托改进的政策，至少可以提高存款人、银行家以及银行借款人的效用。考虑某种替代政策（r，k），它满足 $r > r_P(\bar{k})$ 和 $k < \bar{k}$ 使得银行的特许经营权价值在两种政策之下是相等的（即 $V_P(r, k) = V_P(r_P, \bar{k})$。很明显，银行对于此结果是无差异的，但是存款人的效用提高了。如果在替代政策之下总的银行资产量增加了（即 $D(r, r)(1+k) > D(r_P, r_P)(1+\bar{k})$，那么银行借款人的效用也提高了。尽管此条件并不总是满足的（例如，当储蓄是缺乏弹性的条件下），但是只要总储蓄额的弹性是足够大或者银行资本金的供给弹性是足够小的，那么该结论都成立。

38 还有一方代理人，我们到目前为止都忽略了，那就是银行存款的保险商。模型之中的银行都是同质的，因而监管当局能够实施某种措施来确保均衡条件下不存在任何银行投机行为。当然在异质的银行以及经济之中存在着扰动的情况下，某些银行可能会倒闭。而资本要求可能还会对保护银行存款的保险商起作用。

融自由化倾向于加剧银行之间的竞争程度，但同时银行也被授予了更大的权力自由配置资产组合和决定存款利率水平。其结果就是银行投机的空间也大大增加了。我们一共考察了两种审慎监管的措施：资本要求以及控制存款利率水平。首先，我们指出了在动态经济的情况下，资本要求的作用并不总是像从前想像的那么大。其原因在于它除了能够降低银行投机激励的当期风险资产效应之外，还存在着远期银行特许经营权的价值效应，后者能够增加银行投机的激励。然后我们证明了尽管在资本要求是非常大的情况下，能够实现均衡条件下银行有效地投资，但是这种均衡结果不是帕累托有效的。特别是我们总是可以找到某种结合了资本要求以及存款利率控制的监管体制，而后者总是帕累托占优于任何采用单一的资本要求的政策体制。

我们的模型总共建立了三种经济学直觉：银行存款利率自由决定会损害银行特许经营权的价值；资本要求是有成本的，因为它迫使银行持有更多的昂贵资本；在实施资本要求的时候结合存款利率上限措施，无论后者的约束是紧的还是松弛的，都会大大增加政策的弹性使得政府能够放松对资本量的限制，从而降低资本要求所造成的总成本。

尽管我们着重强调了采用控制存款利率的措施是可以作为避免银行道德风险难题的监管措施的一个补充，但是也必须认识到还可能存在着其他的能够改进资本要求的政策工具。例如资产类型限制以及风险暴露的规则等，都能够帮助解决道德风险难题，特别是能够限制银行参与某些有利于投机的经营行为。存款保险的风险升水也同样能够降低银行投机的激励，尽管有效地实施这样的政策以及恰当地选择分析资产类型对于监管当局而言是很富挑战性的。

对于道德风险难题，我们所提出的另一种政策建议，就是采取“限制增长的速度”。我们的分析过程表明，投机策略是建立在银行快速增长以充分利用人为造成的当期投机资产的高回报率。对于银行的增长速度实施限制的政策就能够有效降低投机策略的收益，从而降低了银行偏离均衡的动机。不过这种政策是可以与其他政策相互替代的，只要银行拥有更好的投资机会或者更低的仲裁费用，银行的增长率也同样地会受到限制。

附录：

引理 1 的证明：

首先，我们证明在资本限制水平充分小的条件下

$(k < \tilde{k}$,其中 $\tilde{k} \equiv \{m_G(r_P,k) = \theta m_P(r_P,k)\})$,$\arg\max_r\{V_G(r,r_P,k)\} > r_P(k)$，并且 $\max_r\{V_G(r, r_P, k)\} > V_G(r_P, r_P, k)$。为此，我们需要证明 $\partial V_G/\partial r_i|_{r_i=r_P(k)} > 0$

$$\partial V_G/\partial r_i = \{m_G(r_P,k)[\partial D(r_P,r_P)/\partial r_i] - \theta D(r_P,r_P)\}/(1-\delta\theta)$$

由决定 $r_P(k)$ 的一阶条件可以得到 $D(r_P, r_P) = m_P(r_P, k)[\partial D(r_P, r_P)/\partial r_i]$

这就意味着

$$\partial V_G/\partial r_i = \{m_G(r_P,k)[\partial D(r_P,r_P)/\partial r_i] - \theta m_P(r_P,k)\}/(1-\delta\theta) > 0$$

其中 $k \leqslant \tilde{k}$。

由 $\bar{k}$ 的定义我们可以得到 $\max_r\{V_G(r, r_P, \bar{k}\} = V_P(r_P, r_P, k)$。因此只要满足 $\bar{k} < \tilde{k}$，以及 $\max_r\{V_G(r, r_P, \bar{K})\} > V_G(r_P, r_P, \bar{k})$，这就意味着 $V_P(r_P, r_P, \bar{k}) > V_G(r_P, r_P, \bar{k})$。由 $\hat{r}(k)$ 的定义我们可以得到 $V_P(\hat{r}, \hat{r}, \bar{k}) = V_G(\hat{r}, \hat{r}, \bar{k})$。因此如果在竞争均衡的条件下采取单一的资本要求，那么银行从投资安全资产中所获得的利润就必须严格大于投机利润水平。但是在银行没有投机行为的条件下，最大的银行存款利率水平是由任意资产的回报率都相等的条件决定的。由于 $\partial V_P/\partial r < 0$，这样银行所提供的存款利率水平一定是严格低于原来水平的（即 $r_P(\bar{k}) < \hat{r}(\bar{k})$）。对于 $\bar{k} \geqslant \tilde{k}$ 的情况，我们知道 $m_G(r_P, \bar{k}) \leqslant \theta m_P(r_P, \bar{k})$。而 $m_G(\hat{r}, \bar{k}) = m_P(\hat{r}, \bar{k})(1-\delta\theta)/(1-\delta)$。由于当 r 变化时，投机资产利润的下降速度要低于安全资产的利润的下降速度（即 $0 > \partial m_G/\partial r > \partial m_P/\partial r$），这样 $r_P(\bar{k}) < \hat{r}(\bar{k})$。因为 $r_P(\underline{k}) = \hat{r}(\underline{k})$ 并且 $\partial r_P/\partial k < 0$，所以一定有 $\bar{k} > \underline{k}$。

命题 6 的证明：

考虑在无投机均衡条件下最大的存款利率水平。令

$$\begin{aligned}\hat{r} &= (1-\delta)(1+k)[\alpha(\hat{r},k) - \theta\gamma]/(1-\theta) + \delta[\alpha(\hat{r},k)\\&\quad(1+k) - \rho(\hat{r},k)k]\\&\equiv f(k,\alpha(\hat{r},k),\rho(\hat{r},k))\end{aligned}$$

这个定义是 $\hat{r}(k)$ 的隐函数，我们对其进行全微分得到：$d\hat{r}/dk = (f_k + f_\alpha\alpha_k + f_\rho\rho_k)/(1 - f_\alpha\alpha_r - f_\rho\rho_r)$。风险资产效应与银行特许经营权价值的效应之和由下式给出：$f_k = (1-\delta)(\alpha-\theta\gamma)/(1-\theta) - \delta(\rho-\alpha)$。利用 $f_\alpha = (1-\delta)(1+k)/(1-\theta) > 0$ 以及 $f_\rho = -\delta k \leqslant 0$，可以立即得到 $d\hat{r}/dk \leqslant f_k$，这里 ρ 和 α 是内生决定的。

参考文献

Akerlof, George A. and Romer, Paul M. "Looting: The Economic Underworld of Bankruptcy for Profit." *Brookings Papers on Economic Activity*, September 1993, (2), pp. 1 – 60.

Besanko, David and Kanatas, George. "The Regulation of Bank Capital: Do Capital Standards Promote Bank Safety?" *Journal of Financial Intermediation*, April 1996, 5 (4), pp. (160 – 83).

Bhattacharya, Sudipto. "Aspects of Monetary and Banking Theory and Moral Hazard." *Journal of Finance*, May 1982, 37 (2), pp. 371 – 84.

Calomiris, Charles W. and Kahn, Charles M. "The Role of Demandable Debt in Structuring Optimal Banking Arrangements." *American Economic Review*, June 1991, 81 (3), pp. 497 – 513.

Caprio, Gerard, Jr. and Klingebiel, Daniela. "Bank Insolvency: Bad Luck, Bad Policy, or Bad Banking?" in Michael Bruno and Boris Pleskovic, eds., *Annual World Bank Conference on Development Economics*, 1996. Washington, DC: World Bank, 1997, pp. 79 – 104.

Caprio, Gerard, Jr. and Summers, Lawrence H. "Finance and Its Reform: Beyond Laissez-Faire," in D. Papadimitriou, ed., *Stability of the financial system*. New York: MacMillan, 1996, pp. 400 – 21.

Chan, Yuk-Shee; Greenbaum, Stuart I. and Thakor, Anjan V. "Is Fairly Priced Deposit Insurance Possible?" *Journal of Finance*, March 1992, 47 (1), pp. 227 – 45.

Cole, Rebel A.; McKenzie, Joseph A. and White, Lawrence J. "Deregulation Gone Awry: Moral Hazard in the Savings and Loan Industry," in A. Cottrell, M. Lawlor, and J. Woo, eds., *The causes and consequences of depository institutions failures*. Boston, MA: Kluwer, 1995. pp. 29 – 73.

Demirgüç-Kunt, Asli and Detragiache, Enrica. "The Determinants of Banking Crises: Evidence from Developed and Developing Countries." Mimeo, World Bank, 1997.

______, "Financial Liberalization and Financial Fragility," in Boris Pleskovic and Joseph E. Stiglitz, eds., *Annual World Bank Conference on Development Economics*, 1997. Washington, DC: World Bank, 1998, pp. 303 – 31.

Demsetz, Rebecca S.; Saidenberg, Marc R. and Strahan, Philip E. "Banks with Something to Lose: The Disciplinary Role of Franchise Value." *Economic Policy Review* (Federal Reserve Bank of New York), October 1996, 2 (2), pp. 1 – 14.

Dewatripont, Mathias and Tirole, Jean. *The prudential regulation of banks*. Cambridge, MA: MIT Press, 1994.

Diamond, Douglas W. "Reputation Acquisition in Debt Markets." *Journal of Political Economy*, August 1989, 97 (4), pp. 828 – 62.

Dinç, Serdar I. "Bank Reputation, Bank Commitment, and the Effects of Competition in Credit Markets." Mimeo, Stanford University, 1997.

Fischer, Klaus P. and Chénard, Martin. "Financial Liberalization Causes Banking Systems Fragility." Centre de Recherche en Économie et Finance Appliqueés, Université Laval (Canada), Working Paper No. 97 – 12, June 1997.

Flannery, Mark J. "Debt Maturity and the Deadweight Cost of Leverage: Optimally Financing

Banking Firms." *American Economic Review*, March 1994, 84 (1), pp. 320 - 31.

Genotte, Gerard and Pyle, David. "Capital Controls and Bank Risk." *Journal of Banking and Finance*, September 1991, 15 (5), pp. 805 - 24.

Gianmarino, Ronald M.; Lewis, Tracy R. and Sappington, David E. M. "An Incentive Approach to Banking Regulation." *Journal of Finance*, September 1993, 48 (4), pp. 1523 - 42.

Gorton, Gary and Winton, Andrew. "Bank Capital Regulation in General Equilibrium." National Bureau of Economic Research (Cambridge, MA) Working Paper No. 5244, 1997.

Hellmann, Thomas F. and Murdock, Kevin C. "Financial Sector Development Policy: The Importance of Reputational Capital and Governance," in R. Sabot and I. Skékely, eds., *Development Strategy and management of the market economy*, Vol. 2. Oxford: Clarendon, 1997, pp. 269 - 323.

Hellmann, Thomas F.; Murdock, Kevin C. and Stiglitz, Joseph E. "Deposit Mobilization Through Financial Restraint," in N. Hermes and R. Lensink, eds., *Financial development and economic growth: Theory and experiences from developing economies.* London: Routledge, 1996, pp. 219 - 46.

______. "Financial Restraint: Toward a New Paradigm," in M. Aoki, M. Okuno-Fujiwara, and H. Kim, eds., *The role of government in East Asian economic development: Comparative institutional analysis.* Oxford: Clarendon, 1997, pp. 163 - 207.

______. "Financial Restraint and the Market Enhancing View," in Y. Hayami and M. Aoki, eds., *The institutional foundations of East Astian economic development.* New York: MacMillan, 1998a, pp. 255 - 79.

______. "Liberalization, Moral Hazard in Banking, and Prudential Regulation: Are Capital Requirements Enough?" Graduate School of Business, Stanford University, Working Paper No. 1466R, 1998b.

Kane, Edward. *The S&L insurance crisis: How did it happen?* Washington, DC: Urban Institute Press, 1989.

Keeley, Michael C. "Deposit Insurance, Risk, and Market Power in Banking." *American Economic Review*, December 1990, 80 (5), pp. 1183 - 200.

Lam, Chun H. and Chen, Andrew H. "Joint Effects of Interest Rate Deregulation and Capital Requirements on Optimal Bank Portfolio Adjustments." *Journal of Finance*, *June* 1985, 45 (2), pp. 563 - 75.

Radelet, Steven and Sachs, Jeffrey D. "The East Asian Financial Crisis: Diagnosis, Remedies, Prospects." *Brookings Papers on Economic Activity*, March 1998, 1, pp. 1 - 74.

Rochet, Jean-Charles. "Capital Requirements and the Behaviour of Commercial Banks." *European Economic Review*, June 1992, 36 (5), pp. 1137 - 70.

Rothschild, Michael and Stiglitz, Joseph E. "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information." *Quarterly Journal of Economics*, November 1976, 90 (4), pp. 619 - 28.

Shoven, John B.; Smart, Scott B. and Waldfogel, Joel. "Real Interest Rates and the Savings and Loan Crisis: The Moral Hazard Premium." *Journal of Economic Perspectives*, Winter 1992, 6 (1), pp. 155 - 67.

Smith, Bruce D. "Private Information, Deposit Interest Rates, and the 'Stability' of the

Banking System." *Journal of Monetary Economics*, June 1984, 14 (3), pp. 293 - 317.

Stiglitz, Joseph E. "Credit Markets and the Control of Capital." *Journal of Money, Credit and Banking*, May 1985, 17 (2), pp. 133 - 52.

______. "Introduction—S&L Bail-Out," in J. R. Barth and R. D. Brumbaugh, eds., *The reform of the Federal Deposit Insurance: Disciplining the government and protecting the taxpayers*. New York: Harper Collins, 1992, pp. 20 - 35.

______. "The Role of the State in Financial Markets," in Michael Bruno and Boris Pleskovic, eds., *Proceedings of the World Bank Annual Conference on Development Economics*, 1993. Washington, DC: World Bank, 1994, pp. 19 - 52.

Suzuki, Yoshio. *The Japanese financial system*. Oxford: Clarendon, 1987.

Tobin, James. "Asian Financial Crisis." *Japan and the World Economy*, July 1998, 10 (3), pp. 351 - 53.

Weisbrod, Steven R.; Lee, Howard and Rojas-Suarez, Liliana. "Bank Risk and the Declining Franchise Value of the Banking Systems in the United States and Japan." International Monetary Fund (Washington, DC) Working Paper No. 92 - 45, June 1992.

信息与私有化

私有化、信息与激励*

这篇论文讨论了在商品和服务的公共生产和私人生产之间进行选择的问题。实际中，两种经营模式都包含了很重要的授权问题，因此在某些方面非常相似。这里讨论的两种模式之间的主要差别在于当政府试图干涉已授权的生产活动时，两种模式所面临的交易成本是不同的。在公有制下的干涉成本往往要比在私有制下少。在公有制下的干涉更为方便；但在私人生产中，不干涉的承诺更具有可信性，这也会带来有益的激励效应。本文给出了私有化基本定理（类似于福利经济学的基本定理），这些定理提出了在哪些情况下政府生产无法优于私人生产，同时本文还评价了这些情况的限制性条件。

一、引言

长期以来，经济学家都非常关心这样一个问题：哪些商品和服务应该由政府提供，而哪些商品和服务应该由私人提供。最近的“私有化”趋势反映了一种论断，即以前的“分配”是不正确的——一些公共部门内的活动在私人部门中进行会更好。本文提出了一个概念性的框架，并在这个框架下分析了几个核心问题。我们特别关注的是，在私有化决策中，激励和不完美

---

* David E. M. Sappington and Joseph E. Stiglitz, “Privatization, Information and Incentives”, Journal of Policy Analysis and Management, 6 (4), 1987, pp. 567 – 582. 这篇文章是为1986年9月在宾夕法尼亚大学举行的公共部门的私有化会议而写的。本文所表达的观点并不代表贝尔通讯研究（Bell Communications Research）的观点。与 Michael Riordan 的讨论有助于形成本文中的许多观点。非常感谢 Janet Pack，Almarin Phillips，Dennis Yao 以及其他一些会议参加者对本文所提的建议。

信息所扮演的角色。

虽然"公共的"和"私人的"标签可能会让人们觉得这是两种非常不同经营模式，但在实际中，国有企业和私人企业的日常经营有很多的相似之处。两种模式都包含重要的责任委托。国会议员和小股东都不能直接控制企业的日常活动，即使原则上企业是在他们的控制之下的。而对公司经营的监管是由代理人或董事会负责的。首席执行官或总裁也被赋予了很大的影响公司的经营的自由决策权。两种所有制下通常都还有其他一些权力层。两种情况下，权力层级的终端是运用本地情况信息进行日常决策并影响公司业绩的经理们。因此，如果我们观察这两种企业的日常功能，国有企业和私人企业在很多方面都非常相似。

在我们看来，国有制和私有制之间的重要区别在于剩余干预权（residual rights of intervention）。在国有企业中，政府保留了一些对已委派好的生产安排进行干预的权力，并且在其认为必要的时候实施一些重大的政策变动。在私人所有制下，债权人（在破产的情况下）和主要的财务权益人（其可以为接手私人公司所需筹集的必要资金进行融资）拥有一些特殊的干预权；但政府的干预权相对于在公有制下却受到了更大的限制。

然而，即使在这篇论文中，公有制和私有制之间的区别也不是完全分明的。例如，在国家级紧急事件（例如，战争）中，政府会干预私人公司的经营，将资源转向能更好地服务于"社会利益"的方面。而且，在金融危机之时，银行和其他债权人常常会向政府所有的公司实施压力。此外，政府会对私人企业进行干预，使其不至于破产。

因此，我们应该关注的似乎应该是和干预有关的交易成本。[1]在国有企业中，政府对已委托的生产安排的直接干预所带来的成本，通常要低于在私有企业中的此成本。就像我们下面所讨论的，这种干预的方便性既会给国有企业带来潜在的收益，也会带来潜在的成本，必须根据国有企业的其他一些固有的优势和劣势来仔细地权衡这些收益和成本。

事实上，有争议的不仅仅是"干预权"问题，还有以下问题：（1）对干预与否的激励；（2）对承诺干预或不干预的能力的限制。对一种组织模式的选择可以被看作是对干预的成本和收益的改动，因此也可以被看作是改

1　干预的交易成本很明显地是同剩余控制权相关的。关于该类权利的更多讨论，请参见 Sanford Grossman and Oliver Hart, " The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration. " *Journal of Political Economy*, 94（1986）: 691 – 719.

变干预的激励机制，部分地替代了承诺能力。

下一部分我们将给出一个关于私有化的基本定理。该定理给出了在什么情况下，所有的政府目标都可以通过设计一个合适的拍卖来达到，该拍卖售出的是生产某种产品和服务的权力。该定理在私有化问题中所起的作用和福利经济学中的一般定理所起的作用非常类似，而且更具有普遍性：它指出在满足了一定的条件时，政府的参与并不能提高私人市场的绩效。对定理中提到的条件的思考，引出了一个“私有化失灵”（类似于“市场失灵”）的问题，即为什么私有化不一定能够达到最期望的结局。非常不幸的是，确定国有化生产是否能够补救这些失灵是一个更难的问题。为了回答该问题，需要用到关于政府行为的理论。虽然我们不具体介绍这一类理论，但我们的一般性的方法会讨论在哪些情况下，政府生产更可能减轻私人生产中所存在的问题。

虽然在下一部分中我们主要讨论对国有企业和私有企业的选择问题，但我们也会指出，在实际中，对组织模式的选择并不总是两者必选其一的。管制一类的制度安排介于两者之间的连续区域内。那样的制度安排同时有国有企业和私人企业的许多好的（和不好）特征。

二、公共生产的范围

众所周知，当满足一定的条件时，对私人产品（即该类产品被一位消费者消费就会排除被其他消费者消费的可能性）而言，私人生产能够达到资源配置的帕累托最优（福利经济学的基本定理具体给出了这些条件，它是关于生产技术和经济中的信息传播的）。如果收入分配并不是所期望的分配方式，那么需要的只是有限的政府干预——采取一次性再分配的形式。例如，当存在外部性、自然垄断[2]或公共品时，就需要政府干预。在公共品的情况下（一位消费者消费某个商品并不排除也有其他消费者消费该商品），如果存在一个帕累托有效的消费水平，则政府必须决定购买的商品量。虽然人们对每种情况下一定程度的干预都是必要的已经达成了共识，但对干预应该采取的形式却有不同的看法。在目前的讨论中，我们集中于一个方面——私人生产和公共生产。

---

2 我们讨论的由自然垄断者生产的商品包括如国防一类的纯公共品和由政府提供的私人品，如教育。关于该类商品要由公共单位来提供的原因的讨论，请参见 Joseph Stiglitz, *The Economics of the Public Sector* (New York: W. W. Norton, 1986)。

政府在不同的生产形式间进行选择时有三个方面的目标：（1）经济效率——政府希望确保是由在生产方面具有相对优势的生产者进行生产，使用的是合适的生产技术和努力水平；（2）公平——政府期望达到某种分配目标；（3）抽租——政府希望能从生产者那里抽到尽可能多的租（利润）（注意，政府对抽租的关注可以被看作是政府对租的分配的关注的一种特殊情况。此处的分配是在“消费者”和“公司”之间）。

我们的核心定理（该定理是“委托—代理”文献中的结论的应用）[3] 指出了在哪些条件下，上述所有目标都能通过拍卖体系“完美地”实现，通过拍卖体系，潜在的生产者为获得某种商品的生产权而竞标。这一结论要成立，要求存在两个或更多的风险中性的公司（代理人），且两个公司关于成本最低的生产技术的信念是对称的。只有被选择的生产者在生产前才知道真实成本（为了简化分析，我们假设该商品的生产是规模报酬递增的，所以只有一个生产者的时候，产业成本是最小的）。政府（委托人）对该商品或服务的产出水平 Q 给出一定的评价 v，用 $v=V(Q)$ 表示。注意，如果我们使不同人群的消费水平明确地出现在政府的评价函数中，则任何配置目标都可以在 $V()$ 反映出来。以上这些要素刻画了所谓的“理想情况”。

下面这个简单的流程确保了政府能够在理想的情况下实现其所有的目标：政府拍卖掉生产权并根据其对产出的估值获得支付 $P(\cdot)$，即 $P(Q)=V(Q)$。换言之，生产决策完全被委派给生产者，而生产者获得的对产出的支付正好等于该产出对政府而言的价值。实施这种形式的补偿机制的结果是提交最高标价的公司（因此也被选为生产商）在结合了实际生产成本的信息后会选择政府最希望得到的产出量。并且如果风险中性的公司一开始对成本的信念是对称的，则竞标的过程将确保生产者不会获得租。因此政府通过将生产委派出去就可以实现最理想的结果，即使它对生产技术一点都不了解。[4] 我们将该结论——当存在设计合理的拍卖机制时，公共生产并不比私人生产更优——称为私有化理论的基本定理。就像上面所提到的，我们将该定理成立时的条件称为理想情况。

---

3 这类文献中一些早期的重要研究关注委托和控制问题，参见 Stephen Ross, “The Economic Theory of Agency: The Principal's Problem,” *American Economic Review*, 63 (1973): 134 - 139; Joseph Stiglitz, “Risk Sharing and Incentives in Sharecropping,” *Review of Economic Studies*, 41 (1974): 219 - 256.

4 Martin Loeb and Wesley Magat, “A Decentralized Method for Utility Regulation,” *Journal of Law and Economics*, 22 (1979): 399 - 404. 这篇论文中提出了这一观点。

拍卖机制看起来似乎非常简单而具有吸引力，因此一个普遍关注的问题是，是否并且如何委派（私有化）各种政府服务的提供。但是我们并没有发现实施该机制成功的一些重要先例。原因可能是理想情况所要求满足的条件不一定同实际条件相符合。

实际中通常无法实现上面给出的理想结果的原因主要有三点。第一，政府从所选择的生产者处抽租常常存在困难；第二，签订合同的成本和对可行合同的制度性限制使得政府在合同设计上的灵活性降低；第三，很可能产生合同实施方面的问题。为了说明这三大类的原因，我们重新具体分析一下理想情况，具体分析其中所包含的一些关键的并且通常不符合现实的条件。

1. 不完全抽租

即使政府可以选择期望成本最低的生产者，并使生产商以最小的实际成本经营，但在实际中，生产商仍常常获得租。当潜在的生产者是风险厌恶的，当对生产权的竞争受到限制时，以及当政府拥有潜在生产者所不知的相关信息时，就会产生这样的租。我们将依次讨论这三点。

风险厌恶。理想情况的一个关键特征是潜在的生产者不是风险厌恶的。该假设确保了政府不需要向公司支付任何的风险溢价，即使公司拥有的有关生产技术的信息可能非常有限，并且由此导致他们最终获得的补偿非常不确定。

当潜在生产者掌握的有关生产技术的信息优于政府，且其对风险的厌恶程度也比政府更高时，会产生一种重要的权衡。将生产任务委派下去的方式是具有优势的，这是因为更具有信息优势的个体（因此也更有能力进行生产）负责生产决策。另一方面，由于生产者承担了风险，还须对他们补偿。如果政府替私人生产商承担了风险，则政府可以获得更多的租。然而，政府承担风险会降低生产商确保生产效率的积极性。如果政府迫使私人生产商自己承担风险，则有可能承担生产任务的不是期望成本最低的生产商，而是风险厌恶程度最低的生产商（当然，承担风险的成本也如其他成本一样存在）。还要注意的是，如果“政府”比私人生产商更能承受风险，则由国有企业来生产会更好一些：如果由私人部门进行生产，则需要向私人部门支付的额外的租可能超过了政府的成本劣势。

对风险的考虑在选择公共还是私人生产形式时是非常重要的，如果私有化是必要的，其在设计拍卖机制时也是非常重要的（例如，军需合同的成本分担条款降低了政府为风险承担而必须支付的补偿金的数额，但也带来了

实实在在的效率成本。此外，关于对石油和天然气勘探开采租约的特权竞标（royalty bidding）和红利竞标（bonus bidding）的比较，有的人认为在特权竞标下，对萃取的烃收取一定的特权费，私人企业承担的风险要小一些）。但是要刻画出两种情况下风险的范围和集中度并非易事。在国有企业中，大部分的风险最后都是由企业的“所有者”承担的。这里的所有者就是整个国民，风险被分散了。对私有企业而言，股票市场或许可以分散大部分经营中产生的风险。如果股票市场可以起到这样的作用，那么我们就看不出在国有企业或私人企业中的“委托人”和“代理人”的相对风险厌恶程度有系统性的差别了。显而易见，如果风险可以大范围地被分散，则两种经营模式都缺乏很强的促使任务被很好完成的激励机制（同组织行为尤其相关的是决策者是否要承担决策后果。如果要承担，则他们更倾向于采取风险厌恶的办法；如果不承担，则激励机制不起作用）。

好的风险市场的缺乏是市场失灵影响私有化决策一个例证。当私人生产商面临的利率高于政府面临的利率时，就出现了另一个例证。对这种市场的不完美性，就像风险市场的缺乏一样，也可以用不完全信息和交易成本来解释。私人生产商违约的可能性抬高了他们的资金成本，而公有制和私有制之间的“财务成本”确实存在差别，因为两者违约的风险是不同的。在石油勘探开采租约中，国家利率和私人利率之间的差别非常重要，因为在红利竞标下，政府只能获得社会总租中的一小部分。出于类似的考虑，有人会认为政府应该拥有自己的建筑物而不是把它们都租出去（除非对私有产权强补偿的管理优势能够建立起来）。实际上，每种情况下的政府都是间接地从私人部门借入资金，而支付的利率都高于国债的利率。

有限竞争。被选择的生产商能够获得租的另外一个原因是在竞标阶段缺乏足够的竞争。在理想情况下，许多的潜在生产者掌握的有关潜在的生产成本的信息是对称的。当只有很少的竞争者参与竞标，且每个生产者对可能的成本的估计不同，则胜出的投标人一般会获得租。为了限制这种租，通常最优的办法是在事后诱导生产商进行低效率的生产，即使竞标者是风险中性的。直观地讲，之所以会发生这样的扭曲，是因为这样的扭曲使得被拍卖的“标的物”对各个竞标者而言更相似，由此可以鼓励竞标更有竞争性。国防和石油再一次提供了很好的例子。在这两种情况下，竞争一般都是有限的（在最近的石油和天然气油田的拍卖中，平均每个油田只有不到两个竞标者）。一些人认为特权竞标要好于红利竞标，即使特许权费用可能会扭曲烃类萃取的方式。该观点认为特权竞标使得政府从竞标中获得的收入大于开采

不足带来的效率方面的损失。[5]

相关的另一种扭曲方式是：对政府而言，最优的做法是，即使自己的期望成本高于更有效率的生产商的成本，在一些情况下也会自己进行生产。这种可能失去生产权的威胁也会诱使潜在的生产者更积极地进行投标。[6]

在一些情况下，只有一家拥有技术能力的公司能够进行生产。当技术为新技术并且市场经验在保持低生产成本时是一个非常关键的因素时，就非常可能出现这种情况。因此，如果在一段时间内只有一家公司是唯一的生产商，则这家公司会比其他的潜在生产商具有明显的优势。当实际上只有一个有实力的竞标者时，它就可以从政府那里抽租。当只有一家公司生产某种特定产品时，就不存在可以用来同该公司的业绩进行比较的自然基准。因此在不存在事前或事后竞争的情况下，控制和抽租问题就会变得非常困难。但要注意的是，在这种情形下，政府生产也面临类似的问题：如果缺乏效率的自然对比，就很难确认政府生产是否是有效率的。而这种情况正好给政府官僚主义者提供机会来增加他们组织的规模（William Niskanen, "Bureaucrats and Politicians," *Journal of Law and Economics*, 18（1975）: 617－643）。

当然，一个产业中的潜在的和实际的生产商的数量对政府而言不一定是外生给定的参数。通过特意将生产转包给不止一家公司，政府可以确保现阶段生产和未来相关生产的竞争性。此外，政府可以创造自己的公司或者对成立新公司进行补贴。用这些方法，政府可以直接影响相应市场的竞争程度，但这明显是有成本的。

具有信息优势的委托人。理想情况的另一个特征是考虑了信息在政府（或者更一般的“委托人”）和潜在生产商（或“代理人”）之间的分布。

5　有关来自特权竞标的潜在收益的讨论请参见 Douglas Reece, "An Analysis of Alternative Bidding Schemes for Leasing Offshore Oil," *Bell Journal of Economics*, 10（1979）: 659－669。对这一现象所建的正式模型请参见 Michael Riordan and David Sappington, "Awarding Monopoly Franchises," *American Economic Review*。

6　这一结论是对 M. Riordan and D. Sappington（1987）所做的研究的自然延伸。对该结论的解释类似于解释为什么拍卖人报出的价格往往是超过其对标的物的估价的最低价格（参见 John Riley and William Samuelson, "Optimal Auction," *American Economic Review*, 71（1981）: 381－392）。类似的直觉使我们发现由成本高一些的第二供货商来负责生产可以有助于形成一个有效的采购机制。（参见 James Anton and Dennis Yao, "Second Sourcing and the Experience Curve: Price Competition in Defense Procurement," *Rand Journal of Economics* 18（1987）: 57－76; and Joel Demski, David Sappington, and Pablo Spiller, "Managing Supplier Swiching," *Rand Journal of Economics* 18（1987）: 77－97。

在理想情况下，代理人比委托人更具有信息优势。然而，当委托人的信息（例如，关于生产任务本身所具有的困难的信息）要比潜在生产商的信息好，则委托人设计的激励机制的性质可能会透露出一些其掌握的私人信息。为了掩藏该信息（例如，不愿让生产商知道该任务实际上是非常困难的，而且有可能会投入大量的努力和资金才能获得成功），委托人可能会背离上面给出的简单机制。这种背离会导致出现不同的生产模式或不同的租的分配。

石油勘探开采租约又是这种情况的例证。政府根据它所拥有的信息计算租约的价值。然后只有当标价相对于估价足够高时，政府才会出租油田。注意，如果私人部门认为政府是信息充分的，并且能够制定一个等于油田租约的期望价值的最低拍卖价，则竞标者认为他们要在竞标中获胜，出价一定太高了。因此他们根本就不会参与竞标。一般而言，政府具有信息优势的事实会对竞标行为产生一种逆向效应（虽然事实上可能并没有如此极端）。

一般地，如果除了在生产过程中直接和生产商进行交流以外，委托人拥有的私人信息很难传递给生产商，那么选择的组织模式可能是最有助于信息交流的模式。在一些重要的情况下，当考虑是否私有化生产时，是否应该把这方面的因素考虑进来还是一个需要进一步讨论的问题。

2. 签约成本和限制

现在，我们来讨论对合同的制度限制和设计合同时的成本是如何使理想情况复杂化的。制度限制包括对生产商负债的限制以及政府的有限承诺。一部会签约成本来自于所有可能困难的预期。

债务限制 。在理想情况下，生产商是有可能蒙受很大损失的。当生产商提交标价时，唯一的要求只是期望盈亏平衡。当成本信息不完全时，事前的成本预测和实际的成本相比可能是太过乐观了。这种情况下，事后生产商会更希望自己没有同政府签订任何的协议。如果对生产商必须承担的损失存在一个下限（例如，破产法的规定，出于政治方面因素的考虑，等等），生产商可能会毁约，而不是提供自己承诺的服务。在这种情况下，就不再能确保得到政府所期望的理想结果了。

我们来讨论一下对国防用品供应商的债务限制所可能导致的结果。无法提供足够的国防产品所导致的社会损失可能会无穷大——远远超过了私人厂商为了确保业绩所做出的任何承诺。因此，失败导致的社会损失使得对失败可以采取的惩罚措施显得微不足道，所以无法用金钱手段诱使私人生产商投

入适当水平的努力来阻止失败。

然而，这并不意味着国有企业就一定是更优的经营模式。失败（或更一般的是服务变差）时能够对政府官员施加的惩罚也不是没有限的。即使在国有企业中，往往生产的许多环节都转包给了私人公司。因此不管是国有还是私人单位提供的服务，都会出现有限债务限制对经营激励机制的削弱效应。

承诺的缺乏。对债务限制的一种解释是，生产商不能承诺一定会实施其一开始同意的合同条款。政府的承诺能力对上面给出的委派机制的可行性也是非常重要的。

为了达到上面给出的通过拍卖机制实现其最期望结果的目的，政府必须能够做出可信的承诺，按照一开始许诺的方式补偿生产商。尤其是，政府必须能够使潜在生产商相信，它一旦选择了一个生产商，就不会对合同进行再商议，并且支付给生产商的（$P(Q)$）将等于产出对政府而言的价值（$V(Q)$）。因此，即使生产商有可能最后会获得极大的利润，政府也必须做出可信承诺，承诺自己不会通过征税将这部分利润抽走。类似地，对于事前大家一致认为有可能发生的损失，政府不能在事后对生产商进行补贴。没有这样的承诺，生产商知道实际上 $V(Q)$ 不是实际的补偿机制，因此，就不能确保实现政府最期望的结果。

政府的承诺能力也会随着所选择的组织模式而变化。在私人所有制下，政府和生产商之间的关系不是那么紧密，因此使得不会使用公共资金补贴生产商损失的承诺更可信。当然，也有一些顶着风险使用公共资金资助快破产企业的案例。克莱斯勒汽车公司的“新股发行”的项目就是美国近年来的一个案例。然而，这样干预的交易成本（包括公众听证会、仔细的法定审查等）是非常高的。不补贴国有企业的承诺往往更难以遵守。法国的超音速运输飞机事件就是这样的一个例子。用引言中的术语来表示，我们这里讨论的问题就是干预的交易成本问题。公有制倾向于降低政府干预的成本，因此在其他条件相同的情况下就使得这样的干预发生的可能性更高。

签约成本 。到目前为止，我们假设将所有可能的情况都写下来，并假设每种情况下的支付和业绩都达成一致是没有成本的。实际上，合同的协商成本可能非常高，许多可能的情况实际上是不可能预见到的。当生产技术非常复杂并且经常变化时，尤其如此，例如军需品行业。意识到这个事实就会发现将生产完全委派下去并不一定合适。相反，为了处理不可预知的偶发事

件，委托人和代理人都愿意不断地进行沟通和再协商。[7]

整个生产决策集会再次影响到这种事后互动的性质。如果政府只选择一个供应商，那么就会在该供应商和其他潜在供应商之间制造出信息不对称。如果存在有效的资本市场并且生产商是风险中性的，则事后互动所能带来的额外租的期望价值是会在初始的竞标价中反映出来的。但是即使在这些严格假设下，也不能保证是在经济上是有效率的：出价最高的公司可能并不是成本最低的生产商，而可能是事后协商能力最强的公司，而这种事后协商本身就可能是无效率的。[8]

一旦引入了连续互动的可能性，就会产生这样一个问题：哪一种组织模式以一种建设性的方式最有助于这种互动？一个关键的问题是让哪些人参与这个互动是最优的。尤其是，国会是否有监管的责任，还是应该该将监管责任交给司法部？公众是否应该有任何直接的发言权？允许公众对生产商的活动有一些持续的直接影响力可能会破坏国会的权威或承诺能力。就像上面所提到的，限制国会的承诺能力，会降低其影响公众最想要结果的能力。因此，需要权衡这方面的损失和公众直接监督国会和生产商活动所能带来的收益。

未知的收益。我们还观察到对委托人（政府）而言，要完全确认其偏好也并非易事。换言之，即使是委托人也可能不知道收益函数 $V(Q)$。之所以会出现这种情况，是因为 $V(Q)$ 是许多委托人的偏好的集合，该集合很难通过政治体系来确认和交流。[9] 另一个不了解 $V(Q)$ 的原因是目标产品是一种新产品，其特性随着技术的发展变化很快。无论是哪种原因，如果收益函数 $V(Q)$ 不是被完全了解的，那么在理想环境中所能实现的简单而完全的分权化生产就不可行了，而且该程序也不能保证实现有效率的生产。

在许多情况下，政府会赋予私人生产商在一定范围内决定支付函数 $P(Q)$的权力。也就是说，公司有有限的自由来设定定价结构。虽然私人公司可能实际上更了解消费者的需求函数，但赋予它们定价的自由会导致潜在的无效率：胜出的竞标者可能并不是最有效率的生产商，而是最知道如何进

7　这些观点是 Oliver Williamson，*Market and Hierarchies: Analysis and Antitrust Implications*，（New York：Free Press，1975）中所述的观点。

8　例如，关于这个问题的讨论请参见 Alvin Roth，*Game Theoretic Models of Bargaining*（London：Cambridge University Press，1987）。

9　这些都是社会选择文献所强调的标准的集体和偏好显示问题。例如，参见 K. Arrow，*Social Choice and Individual Values*（New York：Wiley，1951）。

行价格歧视的公司。因此，只有在有限的一些情况下，公司面临的支付函数才会正好是当政府掌握了公司的私人信息后会实施的支付函数。

政府不完全了解 $V(Q)$ 的原因可能会影响到政府开出正确的政策“处方”。为了说明这点，假设公众知道个人的偏好，但国会并不知道。在这种情况下，生产商和公众之间的直接互动是更好的一种方式，国会的作用应该受到限制。因此应该由私人企业辨别和满足公众的需求。如果生产商之间的竞争是可行的，并且该商品是“普通商品”——因此消费者很容易了解和评价该商品——就最可能出现这种情况。在这种情况下，政府为了保护消费者而对生产商进行仔细地监督就是无必要的。如果消费者可以很容易地监督生产商的活动，并且当某个生产商不能满足他们的需要时，其他的生产商可供他们选择（也就是如果存在足够的事后竞争），那么消费者就可以自己来约束生产商的行为，使生产商生产能够最大化消费者收益函数的商品，不需要依赖政府来做这些。

当然，关于私有化的争论集中在一些由于规模效应的存在而导致只有少数生产商的行业或者该商品是一个公共品的行业。当商品是公共品时，在监督厂商行为时会产生“搭便车”的问题（因为没有一个个体有足够的动力进行监督，即使监督产生的总体社会收益要大于其成本）。当存在规模经济时，由于缺乏其他生产商，就会出现信息方面的问题。结果，消费者进行判断和评价生产商绩效时就会有困难。因此，在这些情况下，消费者用替换供应商的潜在威胁来进行监督就不再那么有效了。

如果是由于消费者不了解自己的偏好而导致他们对政府收益函数 $V(Q)$ 的不完全了解，那么政府直接监督生产商的行为，以确保生产商按照消费者的利益行事，可能就显得非常重要了。因此，最优治理结构要使政府对转包商活动的干预的交易成本相对较小。为了说明这点，我们来看看教育的例子。我们很难严格地说明，度量是什么构成了“教育”，也很难对此制定详细的合同条款。个人常常也不是很确定应该如何来评价教育，而什么是适当教育的标准也是在不断变化的。由于所有这些原因，通过政府来提供教育会有助于政府对教育的监管，因此它也是一种较好的方式。

3. 实施合同时存在的问题

我们现在来讨论实施合同时会产生的一些问题。这里讨论四个问题：测量业绩的问题对资本化企业的能力的限制、多产品的问题以及复杂的控制体系导致的困难。

不完全测量。在理想情况下，生产商的“产出”是可以被所有人无成本

地完全观察到的。在实际中，业绩包含很多成分，而且一些是很难精确测量的。例如，服务质量的所有方面都很难量化，尤其是当不同的消费者会对相同的产品给出不同的评价和对其有不同的感受时。

当不能容易地监督生产商的业绩时，国有和私有企业都可能产生一个严重的控制问题。哪种模式更适合解决这个问题，是一个需要进一步研究的重要问题。我们只是注意到，上面给出的关于不完全了解收益函数 $V(Q)$ 会导致的结果的评论在这里也适用。

关于资本化的讨论。在理想情况下，生产企业的所有权和管理权没有分离。而且，公司的所有者/管理者可以买卖资产。因此，通过将净现金流从企业转移到所有者/管理者手中，就可以轻松地花掉利润。激励机制的特性使得相关人员的偏好可以完全地联结在一起，可以容易地解决积极性问题。实际上，这样的资本化是不可能的，国有和私有企业的管理者都不会根据他们所创造的利润（贴现值）的增加而获得奖励。明白这点，管理者会采取其他的途径来获得他们的租。在国有企业中，获得的租是政治庇护。在国有和私有企业中，经理主管人员获得的补偿是和职位相关的额外补贴。虽然这种形式的补偿可能会有助于激励主管人员和经理，但明显地会产生不效率。例如，当补贴和利润没有完全联系起来时，主管经理在将额外的收入进行再投资以创造未来利润方面的积极性就不大。此外通过前述的方式抽租，企业就会被“放血”，因此，需要随时监督企业活动。

因此，只要将责任委派下去并且所有权和管理权是有效分离的，则国有企业和私有企业中都存在绩效激励的消减问题。在哪种组织模式下问题更严重还是一个公开的难题。答案部分取决于在这两种模式中，对业绩的监管是否有效，竞争是否有效等。

产品加总。在理想情况下，一家公司只生产一种（“一维”）产品。这就使得这种商品的供应决策是和所有其他商品的不相关的。实际上，委派决策并非如此简单。

当要确保多种商品的供应时，为了得到政府最期望的结果，必须至少有两个风险中性的承包人知道整体信息，能够以最低的成本生产出每种想要的产出向量。这样的整体信息在实际中是不可能得到的。没有这样的整体信息，政府也不得不描绘出生产活动之间的边界，尽管并不知道以什么样的方式做最好。结果，就会产生无效率（确实，避免这些问题的唯一办法是政府提出对整个经济的竞标）。

为了说明这个问题，假设政府希望获得两个截然不同而且相当复杂的产

品。政府知道这些产品非常复杂，但并不完全知道它们之间的生产互补性（在电信行业，这一类的情况比较突出）。为了保证这些产品的获得在经济上是有效率的，政府选择了两个代理人，每个代理人都负责监管一种产品的生产和分配。就像在理想情况下一样，即使每个代理人都确保实现它所负责产品的可能的最大消费者剩余，但总剩余也不一定会实现最大值。通常每个单独的生产商都会忽略设计和生产的互补性，但由于政府缺乏对生产技术的整体信息的把握，这种模式并非不合理。[10]

多种产品的情况还存在一个问题。有可能存在这样一种情况：要生产一种产品（例如，军事设备）就必须同时生产另一种产品（例如，有关机密军事情报的信息）。因此，为了控制一种产品（信息）的分配，就可能需要在另一种产品（设备）的生产模式上做出妥协。

等级控制/多个委托人。在理想情况下，委托人（政府）的偏好在本质上就代表了社会福利。这个唯一的代理人也直接负责目标商品的获得。在实际中，等级控制是一个困难得多的问题。困难至少来自于三个方面。第一，有许多的委托人，而这些委托人的偏好可能并不相同。第二，对被委派的生产安排的控制、补偿机制的设计，以及其他类似事务都授权给了一方当事人（例如，国会），而该当事人可能会在一定程度上为自己的私人利益考虑。换言之，理想情况下的委托人实际是一组委托人的代理人，这个被委托的委托人对最终业绩的评价可能会反映出他个人的偏好，以及（或者代替）他所服务的那些委托人的偏好。第三，控制权力通常都会授给不止一个当事人。当有多个被委任的委托人，每个委托人都有不同的授权命令，并且都控制着不同的政策工具时，得到的结果往往都不是帕累托有效的。[11]

当存在多个委托人时，肯定会出现加总问题。除非委托人都有一样的偏好，否则将偏好统一到一个类似于 $V(Q)$ 的一致的社会福利函数中是一件很难的事。更一般地，每个委托人都控制着支付函数的某一部分。因此，如果有几种商品，委托人可能将第 i 种商品的支付函数设为 $P_i(Q_i)$。总的支

10　关于这个问题的其他一些观点，请参见 David Sappington and Joseph Stiglitz, "Information and Regulation," in E. Bailey, Ed., *Public Regulation: New Perspectives on Institutions and Policies*, (New York: Free Press, 1987)。

11　在管制情况下，对这个问题的正式分析，请参见 David Baron, "Noncooperative Regulation of Nonlocalized Externality," *Rand Journal of Economics*, 16 (1985): 553 – 568。更一般性的分析参见 Richard Arnott 和 Joseph Stiglitz, "The Welfare Economics of Moral Hazard," Princeton University discussion paper, 1985。

付函数则是单个支付函数的加总，并且可能同任一一致的社会福利函数 $V(Q)$ 无类似之处。此外，即使构造并指定这个函数给被委派的委托人，也不能确保委托人一定会无私地最大化该函数。如果很难监督到被委派的委托人的行为，并且其私人利益和“社会利益”并不一致（因此会产生同时控制被委派的委托人和生产商的困难），此时不能实现完全地内化社会偏好。[12]

当我们观察其他相关当事人时，另一个相关的问题就凸显出来。不管公众（也就是委托人们）授予了国会（即被委派的委托人）什么其他的权力，国会始终都是拥有几乎不受限制的权力来对公众征税。国会无法做出可信的承诺来约束自己不使用这种权力。因此，公众可能会想到限制国会所能做的转包安排。例如，可能实施工资限制和限制转包商获得的利润，也可能限制可被实施的工程的类型。

显然，这一类的限制会对激励以及业绩有不利的影响。这种影响所带来的社会损失，是否超过了限制被委派生产安排所带来的潜在收益，似乎取决于目标任务的性质。如果任务很常规且易于理解，并且如果生产商付出了足够的努力，业绩的变化就很小，那么对补偿的限制通常不会对业绩产生太大的影响（在这点上，我们想到了垃圾收集任务。）然而，对自身就比较复杂、风险高且要求专家付出有创造性的、十分的努力才能增加“胜利”的可能性的任务，限制对承包商的补偿所带来的不利影响可能就非常显著。当预期的个人收益很小时，可能就无法吸引到最适合的承包商，如果获得的报酬和其表现并不相当，则承包商可能就没有动力去付出努力。

4. 混合型组织模式

前面我们都集中于讨论服务的公共供应和私人供应两种截然不同的模式。实际上，并不一定需要做出二者必居其一的选择。服务的某些部分可由国家提供，而某些部分可由私人提供。根据政府干预所产生的交易成本，管制一类的控制手段可以被视作是介于公有制和私有制之间的控制模式。

首先，我们注意到生产过程是有许多阶段的。购买原材料，制造中间品，将完成品送到消费者手中。在现实中，即使在“公共企业”里，所有的这些活动也并非都由政府来完成。为了说明这点，我们来看看国防的供

12 对于同自利的仲裁者有关的控制问题的正式分析，请参见 Joel Demski and David Sappington, “Hierarchical Regulatory Control,” Bell Communications Research technical memorandum, 1986。

给。虽然最终的产品是由政府向消费者提供的，但许多的投入（例如，武器和设备）是从私人生产商那里获得的。因此，服务的国家供应和私人供应之间的区分并不是非常明显。两者的之间的所谓区分可能仅在于：构成服务的一系列生产活动中，到哪个阶段私人企业就不再参与，而开始由国有企业负责。

而且，即使完全由私人企业负责某种商品的供应，政府的影响也是存在的。例如通过立法来影响产品的内容和分配很常见。此外，政府立法和行政部门的成员会对私人产业发出“施工呼吁”以影响其行为。

当我们发现一些组织模式融合了国家和私人供应两者的特点时，国家供应和私人供应之间的区别就更模糊了。管制就是这样一种模式。被管制行业（例如，公用事业）中的公司往往都是私人所有的。但管制者，也就是政府的代理人，会仔细地观察它们的活动，控制它们的投资回报。

在某种意义上，我们在前面分析过的拍卖也可以被视作是对成为被管制公司的特权的拍卖，而此处对公司的支付由预先规定的评价函数来决定。但典型的被管制行业和前面给出的理想情形之间还存在着重要的差别：（1）在很多情况下，无法拍卖提供某种特殊服务的权利（例如，有线电视）；（2）管制者很少会根据对公司活动的社会评价 $V(Q)$ 来补偿该公司，而是会对定价函数、允许投资回报率、可允许成本施以限制；（3）在我们这里的这个模型中，总承包商和补偿函数（$V(Q)$ 可以被看作是标明日期的产出向量）都是一次性决定的，但在被管制行业中，会周期性地对这些方面进行回顾和修改。

根据上面列出的这些特点还可以推测出管制的基本原理。由于生产的规模经济，被管制行业一般都缺少竞争或竞争有限。因此，对企业的业绩进行即时监督是管制者的职责之一。被管制公司通常还掌握了复杂的生产技术，这样就很难识别生产商的生产能力以及生产过程中所固有的风险。因此，管制者的职责之二是要收集信息以有助于政策决策的制定和限制公司的租。通过信息的搜集，管制者（如果获得了恰当的激励）会逐渐掌握必要的专门知识来对企业活动进行指导，以最大化委托人的利益。当技术不是很复杂时，对专业知识的需求就没有那么明显了。此处关键的一点是：相对不受管制的私人企业，管制降低了某个政府代理人（即掌握了相关信息的管制者）干预企业时的交易成本。

被管制行业中的复杂生产技术也包括生产中使用的特定资产。由于这类资产的其他用途很少，一旦安装就相当于被征用了。因此，为了在一开始的

时候吸引到这类资产，政府必须能够做出可信的承诺，保证这些资产到位后，不会对所有者施以不公平的剥削。管制机制或许有助于政府做出这样的承诺。被管制公司的投入资本要获得“公平的回报率”的要求，至少增加了政府对公司和其持股人进行剥削的交易成本。而管制者拥有的定价权，实际上会使他们拥有很大的勘探权。

管制者另一个重要的作用是促进风险的分担，但同时不会消除对公司行为的激励作用。通过定义“公平的”回报率，如果出现不好的结果是因为公司自己的问题，管制者可以惩罚公司，与此同时，也可以使消费者承担一些由不能预测和避免的坏事件所带来的损失。当然，这种制定回报率的权力也使得管制者剥削公司的可能性增大。[13]

三、总结和结论

我们已经讨论了当某种产品的生产需要在国有企业和私人企业之间进行选择时需要考虑的一些重要问题。我们主要关注由于对生产环境的信息不对称所导致的问题。我们给出了一种理想情况，在这种情况下，私有化是自然的选择，并列出了实际情况和理想情况之间的差别，以及这些差别对私有化问题的影响。

我们也讨论了国有企业和私人企业之间重要的相似之处。例如，两种模式都需要委派出大量的权力。委派是对不完全信息导致的问题——这种问题在两种组织形式下都会出现的——一种自然的对策。在我们看来，两种模式之间的区别在于对已形成的委托生产关系进行干预时的交易成本不同。国有企业中的政府干预往往更易进行，而在私人企业中这种干预就更困难了。然而，这种区别在实际中也只是一种程度上的差别。

在国有企业和私人企业之间进行选择时，很重要的一点是要考虑到干预的期望收益和成本，以及发生干预的概率有多大。进行这种计算时，有两点需要考虑，一是目标任务的复杂程度，二是快速适应不可预测的以外情况的必要性。如果任务非常新且复杂，则更可能出现一些不可预测的情况。如果对这些事件的快速适应是非常关键的（例如，国防中的一些情况），那么易于进行干预来调整活动并控制再协商所持续的时间，可能就相对更重要了，在这种情况下，国有企业可能是一种更好的组织模式。

13 目前关于谁应该承担失败的核能电厂的成本的争论，一方认为是政府想要间接征用，而另一方则认为这是为仔细决策提供相应的激励机制。

当然，如果干预很容易进行，同时对滥用干预权的控制也就困难了，并且会导致低效率。例如，国有企业的经理知道政府会为任何的超额成本提供补贴，就没有动力进行成本控制。

最后，我们想强调一下，当出于不完全信息方面的考虑，而将权利进行委托时，就会出现激励问题，而不管是国有企业还是私人企业都不能完全地解决这一难题。对组织模式的选择仅仅决定了将来对已委托关系进行干预的交易成本，从而会影响到这种干预发生的概率。